

BUILDING A FRAMEWORK FOR OPEN SCIENCE AT THE MNI

October 2015

Authors:

Sarah E. Ali-Khan PhD¹, Liam W. Harris BSc¹, Kendra Levasseur BSc¹ and E. Richard Gold SJD^{1,2}

1. Centre for Intellectual Property Policy (CIPP), Faculty of Law, McGill University, Montreal, Canada

2. Department of Human Genetics, McGill University, Montreal, Canada

RESEARCH OVERVIEW

Background

A growing percentage of researchers, academics, industry and policy-makers has recognized that science uses its informational resources inefficiently, with negative consequences for society, patients and those involved in research and innovation (Birney et al. 2009; Dove and Hubbard 2011). There is growing support among academia and funders for greater openness in science (International Expert Group on Biotechnology, Innovation and Intellectual Property 2008; Chesler and Baker 2010; Dyke and Hubbard 2011; Caulfield et al. 2012; Morgan Jones et al. 2014; Marks 2015; Institute of Medicine 2015). It is in this context that the Montreal Neurological Institute and Hospital (MNI) at McGill University is considering an institution-wide open science policy. This move is both ground-breaking and challenging because the MNI would be a global first-mover in instigating an open science framework in a large, complex and diverse research institution. In the spring of 2015, the MNI asked Professor Gold at the Centre for Intellectual Property Policy (CIPP) at McGill's Faculty of Law to assist in the MNI's policy development process. Throughout this document, we refer to the proposed institution-wide MNI open science policy as the MNI Open Science initiative (OSI), and to open science generally as OS.

Our Mandate

This document presents the results of a consultative engagement exercise with the MNI constituency. This is the first step in the development of an OS framework at the MNI. The primary goal of this research was to explore what OS means to key stakeholders at the MNI and to inform further Institute-wide deliberation. We conducted 21 interviews with a diverse range of

MNI researchers; our objective was to explore their experiences, opinions and concerns regarding OS. Many of them spoke of the important role of the Neuro REB in facilitating OS at the MNI, thus we also conducted an interview with a member of this REB to examine their perspectives on OS. We analyzed this data using qualitative thematic research methods. We then used multidimensional scaling (MDS), cluster analyses and other statistical approaches to deepen, validate and extend our perspectives on the interview data. To preserve anonymity in this document, we refer to the study interviewees as either researchers or interviewees; we specify their research area only if necessary or pertinent. The only exception is the interviewee from the Neuro REB, who agreed to be identified by name to reduce ambiguity in the research reporting.

Findings and conclusions

Our analysis revealed an overall favourable disposition among MNI researchers toward the concept of OS. Many researchers already share research information to a great extent given current constraints. Several interviewees cited the potential of OS for social benefit and increased research efficiency at the MNI. They emphasized, however, that these potential gains are contingent on the way that the MNI shapes policy to support its stakeholders.

Our researchers identified two important concerns that the MNI needs to address to successfully implement an OS. The first is the need for researchers to benefit from their work before broad release of underlying information. The second is the critical importance of safeguarding patient confidentiality and dignity in the course of OS research. In a similar vein, many interviewees stressed that MNI policy should be guided by reference to what will be of greatest benefit to stakeholders, rather than implementing OS 'for its own sake'. We consider this to be consistent with the approach taken by MNI management so far. It nonetheless raises the question of who, exactly, are the relevant stakeholders. An important early step in the process for MNI management should therefore be to clarify which stakeholders carry the greatest weight in its ongoing process.

Our analysis also indicates that many researchers have little understanding of how the MNI is developing its OS framework, and how advanced the initiative is. Some remain concerned that the MNI will impose its selected framework on them without their substantive participation in the process. Others suggested that they have experience, expertise and infrastructure that could be leveraged by the OSI; thus far, they feel these resources remain unrecognized and untapped by

MNI management. Given these findings, we suggest that MNI management communicate the goals and process of an OSI more clearly, and seek greater levels of inclusion of those with experience in open data sharing at the MNI.

We found a significant degree of overlap in the practical concerns related to and preferred boundaries of OS across the MNI constituency clinical and translational scientists, cellular and molecular researchers, imaging/bioinformatics researchers, and cognitive scientists. Nevertheless, there remain significant ideological and conceptual difference between constituencies. This finding suggests that the MNI is most likely to define a common set of principles if it focuses on practical concerns rather than on overall discussion of the role of university research and industry-academic links. A practical set of principles allow for differential implementation in different research areas. Nonetheless, some ideological issues remain in the background, such as in respect of the need for patents and patenting.

There were several key areas in respect of which interviewees called for flexibility in an OSI. In particular, some questioned whether open data sharing is appropriate in the context of industry-academic partnerships. Others expressed concern that the OSI ought not to exclude some level of patenting of certain outputs. We suggest that some of these concerns may be overstated and the result of uncertainty of stakeholder views rather than of reality. Steps should be taken to engage these third party stakeholders, including industry partners, other academic collaborators, donors, funders, McGill management and the McGill Research Office (responsible for technology transfer), the public, patients, the MNI Research Ethics Board (REB) and the REBs of collaborating institutes. This could do much to allay researchers' concerns and maximize the sustainability and effectiveness of an OSI.

Finally, many researchers say they are sharing data extensively. However, many stated that a genuine commitment to the initiative requires additional infrastructure and resources, branding the MNI and creating a clear and explicit policy OS within the framework of a shared mission. Together, these steps should have the effect of formalizing, simplifying, increasing efficiency and efficacy, and expanding currently *ad hoc* and disparate sharing activities. In sum, interviewees believed that the OSI must not impose extra burdens on researchers. Rather, the policy must make sharing easier and not put them at a competitive disadvantage. Ideally it should be structured so as to be advantageous for their research activities and career development.

INTRODUCTION

Background on Open Science

There is general consensus among researchers, funding agencies and journals in recent years that faster and more extensive sharing of data, samples and other scientific materials is critically important to scientific progress (Birney et al. 2009; Dove and Hubbard 2011). Building on the success of software's 'open access' model (July 2010), scientists and policy-makers increasingly stress the importance of accessibility of scientific resources by encouraging the creation and use of repositories and other distribution channels (Birney et al. 2009; Dove and Hubbard 2011; Dove et al 2012), while discouraging intellectual property protection that could restrict use of those resources (July 2010; OECD 2011; Dove et al 2012; Caulfield et al. 2012; Morgan et al 2014). Many researchers and policy-makers view this so-called 'Open Science' model as a way to encourage accountability and replication of scientific findings, to accelerate new and diverse research, and to decrease the costs of obtaining data, samples and other research materials (Birney et al 2009; Rodriguez et al 2009; Chesler and Baker 2010; Dyke and Hubbard 2011; Dove et al 2012). It is in this context that the Montreal Neurological Institute and Hospital (MNI) at McGill University has expressed interest in developing an institution-wide Open Science policy.

Thus far, we have seen Open Science policies adopted in respect of specific large-scale projects rather than at an institution-wide level. These projects are often, but not exclusively, undertaken in the public interest, and find encouragement by journals and funding agencies (Dyke and Hubbard 2011; Caulfield et al 2012). More comprehensive, institution-driven policy is largely restricted to institutions with defined research agendas in fields already familiar with large-scale open data sharing (Dyke and Hubbard 2011). Implementing Open Science in a large, diverse and complex research institute like the MNI presents a much greater challenge.

MNI management approached Professor Gold's team at the Centre for Intellectual Property Policy (CIPP) at McGill University's Faculty of Law early in 2015 to request assistance in the development of their Open Science policy through rigorous and transparent methods. This report represents the first step in this process – the consultative engagement of MNI researchers through in-depth, face-to-face interviews aimed to explore their opinions, concerns and experiences regarding Open Science. The MNI management plans on using these results as a starting point for dialogue with MNI researchers to develop guiding principles, best

practices and detailed protocols to support the introduction of an Open Science framework at the MNI.

Project goals

The primary goal of this research is to explore what Open Science means from the perspective of the MNI constituency, as a first step toward developing an institutional Open Science policy that is sensitive to stakeholder values, and thus more likely to be effective and sustainable. Our interviews aimed at examining the breadth and boundaries of Open Science, including but not limited to the following questions: what types of data and scientific resources may be shared; time frames for sharing, caveats and conditions for sharing; how the conduct of science, including with third parties or involving research participants may be affected by an obligation to freely share data and other resources; the potential role of intellectual property rights; and other benefits, challenges, concerns and potential solutions. We also gauged the potential implications of Open Science policy for the MNI's research, educational and associated activities.

A second and important aim of the CIPP research is to document and publish in respect of the MNI's research and policy development process so that it may be shared with potential funders and other institutions with an interest in the Open Science model. Greater openness in research and innovation is a central and increasingly important factor shaping global scientific and technological development (Chesler and Baker 2010), yet cultural norms and practices remain nascent. Ultimately, we hope this work will contribute to the development of best practices and cultural norms not only for open science implementation at the MNI, but also beyond.

In this document, we refer to the proposed institution-wide MNI Open Science policy as the MNI Open Science initiative (OSI), and to open science more generally as OS. To preserve anonymity, we refer to the study interviewees as either researchers or interviewees; or where necessary, we also specify their research area.

RESULTS

Overview

We interviewed 21 MNI researchers and one representative of the Neuro Research Ethics Board (REB) between June and mid-September 2015 (see Table 1). We analyzed this data in

two steps. First, we used qualitative thematic analysis techniques to identify opinions, experiences and concerns regarding the OSI and OS more generally. Our analysis yielded 10 major themes of discussion, which we present below with their sub-themes. They are:

1. Response to and knowledge of the MNI OSI
2. Understanding the definition of OS
3. Motivation to participate in OS
4. Scope and comfort with sharing of scientific resources
5. Industry partnerships under an OSI
6. Intellectual property
7. Attribution and publication: 'staying ahead'
8. Infrastructure, financial, human resource and time allocation
9. Human samples/data and REB approval
10. Public trust and consent

We sought to provide a visual and over-arching perspective on the differences and similarities in the viewpoints expressed by researchers, and to identify key issues in respect of which there are significant differences of opinion within the MNI constituency. We also explored how views differ along demographic lines, such as gender and research focus. This knowledge may assist MNI management to better understand sub-groups within the Institute, and to more effectively engage and address concerns during policy development. We present these analyses after the thematic analysis (see Appendix 1 for a detailed description of research methods).

Table 1: Breakdown of study interviewees by research focus

Research Area	Number interviewed
Clinical and translational	6
Cognitive neuroscience	2
Imaging and bioinformatics	5
Neuro REB	1
Research managers (included in tally above)	2
Post-docs (included in tally above)	1
Total	22

Themes

1) Response to and knowledge of the MNI Open Science Initiative

Overall, researchers demonstrated a willingness to engage in the interview process and to share their viewpoints, enthusiasm and concerns. We explained during the interviews that this was the first step in a consultative engagement process that will lead to further dialogue between researchers and MNI management. Interviewees welcomed this information. The majority expressed positive opinions toward the ethos and goals of OS in theory. However, almost all voiced some concern or misgiving, which may deter their full participation in a proposed OSI.

Uncertainty about the MNI OSI

Interviewees expressed varying degrees of awareness and understanding of the MNI OSI. The majority, including those who had attended the town hall event, were uncertain about what OS would entail at the MNI, the implementation process followed by MNI management, and how far this process had progressed to date. Only a few researchers reported feeling informed about the initiative. Several interviewees were enthused by Aled Edwards (of the Structural Genomics Consortium at the University of Toronto) and his description of OS as a mechanism to boost research translation, attract funding and accelerate the expansion of scientific knowledge. Yet most interviewees were unsure about how this would work at the MNI. For example, some basic researchers noted that the SGC is calibrated for drug discovery, a focus of little interest to them.

A need for broader researcher engagement and consultation on OS

Some researchers expressed concern that the MNI would impose mandatory requirements to share resources or forgo patents without their substantive consultation on this policy. Further, a few researchers expressed doubt that MNI management has a realistic understanding of the practical challenges in launching a comprehensive and sustainable OSI. Interviewees who had experience with the banking and sharing of patient samples underlined the considerable time, economic and human resources involved. Some said they felt that management is advancing the OSI without due consideration of the potential adverse effects on their work. Researchers involved in clinical and translational research expressed, in particular, this view. The majority of interviewees pointed to a need for further consultation and education with researchers to promote comfort and encourage participation. Notably, some said that postdoctoral and non-

tenured research staff, including research managers, should be included in consultations because it is often these individuals who facilitate the practical implementation of OS.

Untapped MNI resources

Some interviewees noted that substantial knowledge and infrastructure for implementation of OS already exist within MNI research groups. Bioinformatics and imaging researchers underlined their extensive experience and existing infrastructure for data-sharing, which they said has thus far been under-recognized by MNI management.

Points to consider:

- Some researchers felt that an OSI would not represent a major shift in their current practices, saying they are already participating in OS sharing of their scientific resources to a significant extent.
- A good number of interviewees said that an OSI that is imposed on researchers without in-depth consultation and without allowing for flexibility in its implementation would not be widely adopted by the MNI constituency.

“I speak as someone who is very much in support of open science, and I want to see it happen, but I’d rather use a carrot than a stick [...] I wouldn’t go so far as to make it a condition of employment at the MNI, for instance.”

- Clinical and translational researcher

2) Understanding and definition of OS

When asked for their personal definition of OS, many researchers identified certain key features: data sharing (some specified pre-publication sharing); reagent sharing; Open Access publication; forgoing patents or relegating them to more applied research outputs; informal sharing of expertise and knowledge through teaching, mentoring and oral academic or clinical exchange; and greater transparency and access to information for patients. Few researchers were aware that OS can extend to avoiding the use of research tools or reagents covered by the restrictive (such as reach through rights) intellectual property of others. Some interviewees noted that the definition of OS is potentially very open-ended, and that establishing a clear definition is critical to developing an effective OSI and earning researcher buy-in.

“So I think it has to come to the mandate of the institute... It’s not just to release stuff for the sake of it”.

- Basic scientist

Points to consider:

- Several interviewees stated that OS policy should be grounded within MNI mission and vision statements. They noted that there is no clear and recent institutional statement, and that articulating an explicit statement should be a first and fundamental step toward developing the OSI.
- Ultimately, many emphasized that the definition of OS at the MNI must be guided by what will maximize benefits for stakeholders. They listed a variety of potential stakeholders, including researchers, the public, patients, funders, donors, and industry and academic research partners.

3) Why participate in OS

Researchers described several motivations for the sharing of scientific resources they already undertake, and by extension for participation in OS. Some of these motivations stem from moral and others from practical grounds.

Motivations for sharing and greater openness

Almost all interviewees emphasized that sharing, transparency, and openness with data and resources is a central tenet of the academic model. They said that it is good practice to share, such that a community norm is strengthened and that others will do the same. Some further specified that although responding to requests for reagents and data can be ‘a pain in the ass’, they do so in the hopes that others will reciprocate when they need it. A number noted that sharing has other professional benefits. Some said that their sharing activities have increased their professional stature and their citation rate. Many said that requests for reagents or the use of shared data, algorithms, software or platforms has led to new and diverse collaborations allowing them to expand their interests and their professional impact. Some said that sharing leads to more publications, and they expected that adoption of an OSI would enhance researchers’ publication rate.

Several interviewees pointed to an ethical impetus to share, saying that it is ‘just the right thing to do’ or noting that work product resulting from public funding should rightly be placed in the public domain. Many spoke of social benefits or of contributing to the ‘greater good’, saying that sharing potentially maximizes the value of findings or data by allowing others to use and benefit from them. Some articulated a related economic argument: that broad sharing optimizes the value of investments in research.

“What typically happens in the past, a lot of studies would occur and then... that data was effectively lost. It was used for a very specific purpose and then, you know, years later it’s gone. No one has it. So, why not, I mean we spent millions of dollars to acquire this data, why not have it for the rest of time, available to a wide variety of people?”

- Research manager, imaging and informatics

Several interviewees, including basic, and clinical and translational researchers, spoke about patient data. They felt that it is unethical for researchers to withhold this data from the public domain (subject to ethics requirements and informed consent) because the data belongs to the patients. They reiterated that patients are often motivated to participate in research to ‘find cures faster’. As such, making patient data accessible to the greatest number of researchers is an important step toward accelerated innovation. Similarly, one researcher noted that high throughput datasets should also be posted pre-publication such that others can make use of these valuable and extensive data. Many researchers referred to the volume of data ‘bottlenecking’ before publication; this information could be put to good use more quickly. The majority of researchers mentioned that increasing accessibility to data and findings more generally allows ‘many brains to work on the same problem’, thus augmenting the efficiency of research and the advance of knowledge. Thus, they said that data can be validated and replicated earlier, and duplication reduced. Some also mentioned equity arguments, noting that those in lower resource settings would be able to access knowledge more readily if science were more open. Finally, one researcher said that sharing puts information in the public domain, preventing others from unauthorized commercialization of her/his work.

Motivations for reducing IP (see Theme 5 and Theme 6 for more details)

A number of researchers expressed disdain for claiming ownership over data or knowledge that has public value and would be of broad benefit if placed in the public domain. Some held that IP

can impede research progress; some cited examples from their own experience. A few said that the OSI would increase the efficiency of research at the MNI, by reducing 'red tape' – time, money and human resources spent on pursuing IP that turns out to be worthless. A few specified that industry would be more attracted to collaborate with MNI researchers if access to knowledge was less restricted by IP.

Stature of the MNI (see Theme 10 for more details)

A number of interviewees said that the OSI could increase the stature of the MNI, by branding the Institute as a leader in the field. Some noted that depending on the how the OSI is presented to the public, it could contribute to perception of the MNI as a place where research translation and therapeutic innovation is optimized for the greatest social benefit. Thus, greater transparency, openness and less IP could underline the perception that research benefits flow more directly to patients, and are not simply appropriated by industry. Such perceptions could enhance public trust of the MNI and may lead to greater patient participation in research, inspire more donations, and attract researchers and trainees interested in OS to the MNI.

Points to consider

- Many interviewees emphasized that safeguarding patient confidentiality and dignity is critical in the conduct of OS. They specified that that primacy of this concern must be front and centre in presenting the OSI to patients and the public.

4) Scope of comfort with the sharing of scientific resources

Overall, interviewees said they already share a range of scientific materials, though they distinguished between sharing with established collaborators/co-grantees and broader sharing with the international scientific community. Almost all interviewees reported sharing data on *ad hoc* basis with collaborators, others at the MNI, or when resources are requested by outside scientists as standard practice. Some noted that they share full datasets underlying published findings as supplemental online files, as required by many journals. Few outside the bioinformatics and imaging area said they take part in larger scale or open sharing activities such as contributing to public repositories or developing their own open access databases or biobanks. An exception is that a number of researchers use public repositories such as

Addgene and Jackson Labs to achieve wide distribution of plasmids and knock-out models respectively.

Qualifiers for sharing: a spectrum of appropriateness

Researchers described a variety of qualifiers or caveats to sharing scientific resources. The majority were very clear that sharing must be done in a just and sustainable manner given their competitive and academic work environment (see Theme 7: Attribution and publication for further details). Thus, interviewees widely stated that experimental data should not be released pre-publication, because this could lead to 'being scooped,' or preclude development of a more comprehensive and high impact scientific story, thus harming funding and career advancement.

"the concept to me is, if you work hard for something, you should have the privilege of using it. ... I've been in situations where I have worked very hard to collect data, to obtain the data that I think are interesting, I don't want to just make this data available to everybody when I finish collecting it"

- Researcher

There was almost unanimous agreement among interviewees that it is appropriate to wait until after publication to share experimental data beyond established collaborators. Most stated that this is already their practice. Most interviewees stated a preference to fully understand and publish before publicly releasing a dataset or reagent. Further, a good number held that it is socially irresponsible to release data before it is fully validated and its quality is assured, as this could waste others' time and promote erroneous conclusions. Several noted the potential for openly shared data to be used in research or for applications with which they do not agree. They worried that their own or the MNI's reputation might be compromised by such associations.

In general, when researchers invested more time or intellectual and economic resources in the generation of reagents or data, they expressed a greater desire for control over when, under what conditions and how widely those resources could be shared. For example, a few specified that it is unethical to delay the broad sharing of clinical and genomic data collected from patients, (in the case that ethical requirements are met and informed consent is obtained) because these entail minimal intellectual input from researchers and belong to patients. At the other end of the spectrum, researchers who went through a costly or lengthy process to generate a resource -- for example iPS cell-lines or certain knock-out mouse models -- felt that it is only fair that they

extract full research and publication value before sharing with the broader community. Likewise, researchers who invest significant time and energy in maintaining depletable or precious biosamples cited the need for tight control over access. Thus, in the case of high-value reagents, researchers said they carefully review requests and will not share if the user wishes to pursue scientific questions that are too closely related to their own. Further, some said they prefer to establish ongoing collaborations with those who seek access, rather than release reagents with minimal further contact. Thus, many interviewees distinguished between different types of scientific resources in regard to the ease and ethical implications of sharing, and their willingness to share. The imaging and bioinformatics field appeared to an outlier in terms of the extent of open sharing already taking place. Researchers in that field said that sharing data and other resources including software, platforms and algorithms with the international scientific community is already the norm in their field. Nevertheless, even those researchers, whom we expected to be the most comfortable with open pre-publication sharing, largely favoured releasing resources after publication.

Table 2: Examples of scientific resources and issues that impinge on the ease or difficulty of sharing, or willingness to share mentioned by interviewees.

Examples of scientific resources	Issues impinging on the ease or difficulty of sharing, or researchers' willingness to share
All resources	In general, there is agreement that broad sharing should be reserved until after publication, and sometimes delayed until researchers are satisfied that publication value has been extracted.
Patient data, including genomic and clinical data	Sharing is subject to REB approval and patient preferences as specified in consent forms; Patient data must be anonymized before sharing; It should be recognized that this data belongs to the patient, thus if she/he wishes it to be shared, it is unethical for the researcher who collected it to withhold it for her/his exclusive use.
Patient bio-samples	Sharing is subject to REB approval and patient preferences as specified in consent forms; Associated meta-data must be anonymized before sharing.
Depletable or precious biosamples	Researchers largely prefer to reserve use to themselves and close collaborators; Researchers wish to review applications for access; Researcher often prefer to enter into ongoing collaborations with those with whom they share these resources.
Imaging data	The format of this data and the existence of substantial MNI and broader community infrastructure sometimes allow this data to be readily shared; Lack of interoperability of data types and platforms, and data storage capacities can be a barrier to sharing; Sharing is subject to REB approval and patient preferences as specified in consent forms; Identifying features (such as representation of the patients face) and meta-data that

	may compromise patient confidentiality must be removed before sharing; Obtaining REB approval to share (and use) data in Web repositories are often time and resource intensive.
Cognitive/behavioural data	There is a lack of standardized taxonomies and formats to allow large-scale sharing: developing these is likely to promote sharing, but would be a very significant undertaking.
Plasmids	Public repositories exist that can administer sharing, thus divesting researchers of the ongoing time and cost involved.
Antibodies	Researchers said there are no existing public repositories: thus labs must administer sharing themselves - this can be time and resource intensive. Antibodies can be licensed to commercial enterprises, thus enabling broad access to a standardized product, but for a price.
KO mice	Public repositories exist that can administer sharing, thus divesting researchers of the ongoing time and cost involved; In some cases researchers would like to extract full publication benefit before sharing.
iPS cell-lines	Substantial cost and time investment is involved in generation – generally, researchers would like to extract full publication benefit before sharing; Care must be taken with sharing any meta-data that may compromise patient confidentiality – however, these are often key information for ongoing research.
Algorithms/ software	Can be readily shared through the Web, and click-wrap and copyleft licenses can be used to ensure these remain Open Source.

Pre-publication sharing

Many interviewees noted that sharing pre-publication data in a conference setting is an important standard academic activity. While most said that they currently do this, some specified they do so judiciously either to prevent ‘being scooped’ or to allow for eventual patenting. A few of the more established researchers noted that greater protectiveness over experimental data is appropriate for those at earlier stages of their careers. Thus, these interviewees said they take care not to jeopardize their trainees’ research when presenting publically.

Disincentives and barriers to sharing

While many researchers said they aim to broadly share data and reagents post-publication, they noted that there are a number of barriers and disincentives in practice. The most frequently mentioned issues according to resource type are shown in Table 2 and discussed in more detail in Themes 4, 5, 7, 8, 9 and 10. Some noted that another reason why they mainly share reagents post-publication is that publication is the key means by which potential users learn of the existence of a reagent.

Table 3: Examples of some of specific concerns and disincentives to participating in an OSI related to a selection of resource types that were most mentioned by interviewees. Concerns are listed with the issues mentioned most frequently from top to bottom.

Imaging Data	Clinical Data	Reagents	Compounds	Cell lines	Animal Models	Software and Technology
Time, financial, human resource burden	Ethics approval processes are a barrier	Time, financial, human resource burden	Risk of losing collaborations because these require secrecy about data generated	Significant researcher time and money invested in resource creation – reluctance to share	Time, financial, human resource burden	Commercial entities might take advantage of openness
Concerns around privacy/confidentiality of research participants	Concerns around privacy/confidentiality of research participants	Shared resource may be used out of context/un-informed users will lead to poor research outcomes/concerns about sharing of un-validated/poor data	Risk of losing private investment in research if there is a requirement for data-sharing and absence of IP	Commercial entities might take advantage of openness	Shared resource may be used out of context/un-informed users will lead to poor research outcomes/concerns about sharing of un-validated/poor data	Shared resource may be used out of context/un-informed users will lead to poor research outcomes/concerns about sharing of un-validated/poor data
Ethics approval processes are a barrier	Losing competitive advantage by sharing too much			Losing advantage by sharing too much		
Lack of interoperability of data, software and platforms between potential sharers				Time, financial, human resource burden		
Insufficient infrastructure for sharing						

Additional points to consider:

- Some suggested that requiring registration to access shared data would be helpful in mitigating social costs from potentially under-informed or irresponsible users. A few suggested that access requests be reviewed by a committee and/or by the individual researchers who shared the resource or data.

- Interviewees who discussed high cost, scarce or depletable resources agreed that a governance mechanism to access these materials is essential.
- At least one interviewee questioned whether the OSI could rightly be called ‘open’ if access to resources requires registration or ‘vetting’.

5) Industry partnerships under an OSI

Our findings suggest that one of the greatest challenges to an OSI may be sharing data generated through research relationships with commercial partners. Researchers said these relationships are desirable because they provide valuable research funding and scientific opportunities. Additionally, researchers emphasized that these relationships provide important avenues for clinical translation and patient benefit -- goals that many said are aligned with the mission of the MNI. They mentioned two common paradigms involving the pharmaceutical industry: the screening of compound libraries in assays developed by MNI researchers and the clinical testing of compounds.

Most researchers involved in these types of relationships said that commercial partners are unlikely to allow them to share resulting data openly. Around half of the interviewees specified the potential loss of this kind of private investment as a key disincentive to participating in OS at the MNI (see Table 3). However, a number of researchers conceded that they had not yet discussed the possibility of greater openness with partners nor their partners’ attitudes to OS. Other researchers said that greater openness is likely to attract industry partners, because the absence of IP and ready access to data would streamline moving discoveries to commercialization. When discussing this possibility, researchers often referred to Aled Edwards’ discussion of OS. A few interviewees said they thought industry partners would seek patents over outcomes, something that is not necessarily inconsistent with OS provided that the MNI is not involved. Further, some noted that the need to explain and approve OS-related protocol changes with collaborators and collaborators’ REBs would likely extend project timelines and that these delays may not be acceptable to industry partners.

Points to consider:

- A few interviewees recommended developing a limited number of formal “models of engagement” to allow for flexibility in the structure of interactions with external partners under an OSI. Again, they underlined that the specifics of an OSI (i.e. whether to engage with an industry partner, whether to share specific samples) must be determined in

terms of what is most beneficial to stakeholders, especially patients and MNI researchers.

- More than half of interviewees suggested that an OSI should allow a degree of flexibility or the option to opt-out. For example, several advocated for the separation of the Clinical Research Unit from the OSI. However, a few pointed out that switching to a closed model when OS becomes inconvenient would challenge the integrity of the initiative.

“I would think of myself as someone who is a proponent of the open access concept and philosophy - at the same time also recognizing the complexities... if engagement in open access here was an all or none ... that will have more people have to decide not to participate than otherwise might participate for at least some of their activities”.

- Clinical and translation researcher

6) Intellectual property

Interviewees expressed the greatest uncertainty about OS at the MNI on the issue of intellectual property (IP). We heard a range of viewpoints concerning the appropriateness of patents in MNI research.

MNI research product is ill-suited for IP

Some interviewees, in particular basic and clinician-scientists, consider their research product to be ill-suited for patent protection. For example, a few pointed out that knock-out mice cannot be patented in Canada, or mentioned the recent US Supreme Court decision in *AMP v Myriad* as an important signal that patents on basic scientific discoveries, including genes and receptors, are inappropriate. As such, a good number of interviewees across all research settings noted that patents are often pursued too early in the research process or on inappropriate entities.

“I think that at the level of the kind of academic research we’re doing, a lot of our findings may get protected by intellectual property - which will tie things down and slow down the process, when only a really infinitesimal fraction of those findings ever really make it to any kind of real commercial value, or are ever translated in to something that will be commercially, you know, profitable or generating revenue”.

- Clinical and translational researcher

Disinterest or lack of time to pursue IP

Many interviewees expressed little interest in pursuing IP. A good number said they do not wish to waste time and money on these activities because patents are rarely fruitful from either a commercial or translational standpoint. Many noted that their primary motivation is academic discovery, not commercial gain. Several interviewees stated that they do not have the entrepreneurial drive or expertise to make patenting of discoveries worthwhile, while others simply were not interested in holding ownership rights over their research product. Further, one researcher pointed out that patents must be enforced to be effective, noting that researchers generally lack the time for such surveillance.

Several researchers expressed disdain for pursuing commercial gains from their research. Some held that publicly funded scientists should be working collectively to advance science in the public interest, not for material benefit. A few researchers have established companies in which key technologies are not patented. They said the success of these ventures relies on staying one step ahead of competitors by pushing advances in the field and the expertise and tacit knowledge of those involved, rather than ownership of key inventions. Finally, a few noted that they use some forms of IP as a defensive mechanism, for example to discourage commercial entities from appropriating software or experimental protocols.

A right to patent

Some interviewees said they hold patents or employ other forms of IP such as copyrights and licences. One said she/he licenses antibodies and protocols to make reagents to companies – this relieves the lab team of the resource cost of distribution and makes these reagents widely available in a standardized fashion. Additionally, it is a welcome source of revenue. Several interviewees stated unequivocally that the right to claim IP protection is an important aspect of their academic freedom, and a justifiable manifestation of their creativity. A few cited the potential for personal material benefit as a motivation to patent.

Patents as mechanism to facilitate translation

Despite the number of interviewees who lacked interest in holding IP, there was broad agreement that patents or other modes of protection on research products are critical in raising funds for translation. Many emphasized that translation to patient benefit is the essential goal of the MNI's collective research activity.

“Socially it is nice, other scientists can use your stuff and some companies may develop your ideas because it is open and it is available, but very often it requires investment for things to become ... available to all the patients in the world... And so if there is no patent it's nice, it's open to the world, but it may never be developed because it's not protected. So that would be terribly negative, a very negative effect.”

- Clinician-scientist

Implications of the IP protections of others

A key tenet of OS is that research products should be readily available to be replicated and further developed by the scientific community. Protections held by others on research tools can limit the ability of downstream users to reproduce or use the results. To address this issue, we asked researchers to what extent they use patented resources, or those subject to other restrictive protections. Many reported using patented material in their research, and some stated that using non-patented materials would actually hinder their research. Others said that although they currently use patented inventions, they could be avoided. Some interviewees said that the IP protections of others -- including copyrights, material transfer agreements (MTAs), reach-through licenses and patents -- had slowed their research progress. Often interviewees framed MTAs as excessive and unnecessary red tape, but allowed that the materials gained were useful enough to justify the time spent. Few admitted to asking others to sign MTAs.

Points to consider:

- Some interviewees who currently hold IP on aspects of their research output said they would be willing to consider forgoing patents in the future. This is largely because the patents they currently hold have rarely turned out to be worth the time and money spent.
- Ultimately, most researchers who discussed patents agreed that a clear provision allowing IP (although, perhaps not by the MNI or its researchers) in specific and beneficial circumstances would be an essential component of an OSI at the MNI. However, they were uncertain about the appropriate mechanism for defining such circumstances.
- A few suggested instituting an internal and confidential MNI patent review committee, to decide which patents are worth pursuing.
- Others suggested that if the MNI was to become a patent-free zone, IP on worthy inventions could be held by external linker entities. However, researchers were unsure of the details of how such a mechanism would work.

- In some cases it was clear that researchers themselves would want to be owners and beneficiaries of any patents pursued.

7) Attribution and publication: 'staying ahead'

Almost every interviewee mentioned that they are operating within a competitive research and academic paradigm in which publishing papers in high impact journals and accessing funding is increasingly difficult. Every interviewee, except the few involved exclusively in clinical research, stated that their key priority was to maintain strong metrics to support funding applications and academic advancement.

Fear of losing a competitive advantage

The majority of interviewees voiced the concern that participation in an OSI had the potential to weaken their competitive advantage if researchers had to release resources (data, reagents, biosamples etc.) before the researcher was able to extract value and before publications and/or patents marked the researcher's paternity of those resources. Thus, researchers said that pre-publication release of data or reagents might preclude publication in top tier journals by reducing the novelty of findings, or might lead to being 'scooped'. Both of these outcomes would have adverse consequences on the researcher's ability to secure future funding. Some noted that these issues are more significant for early-stage researchers, while others noted that the advantages of OS might mitigate negative effects (See Theme 3: Motivations to participate in OS).

"I would like to be more open as, as open as possible, you know. Because I would like the findings to be distributed, diffused and people will work on that, there will be a follow-up and so on - that's perfect. But in the system we are working in, we have to publish in the best journal and for that we have to be competitive, for that we have to be secret with some confidentiality linked to that. So that's where we lose the openness"

- Molecular and cellular researcher

Some interviewees were concerned that open sharing of pre-publication findings or reagents would allow commercial entities to appropriate their work without compensation or acknowledgement. Others noted that making data and other resources openly and freely available could remove a key motivation for collaboration with the MNI. For example, several noted that the MNI's unique bio-sample resources attract high quality collaborators (both

industry, and other academic and publicly-funded partners). Allowing open access might remove this motivation for collaboration with the MNI.

“We have something that nobody else has. They want it, it has value.... So if you were open, does open mean you give them everything? Then yeah, you would have lost the opportunity for a \$25 million grant”

- Clinician-scientist

Finally, a few outlined a worst case scenario in which an OSI might diminish the value of the MNI itself because the most interesting findings would simply be taken over by others, including other academic groups. A number of interviewees worried that the MNI would go too far ‘out on a limb’ with the adoption of OS before other institutions in Montreal, in Canada and beyond. As such, others might access MNI resources without reciprocating. Thus, while a good number of interviewees said the culture of science is undergoing a shift toward greater openness, many cautioned that the MNI should advance with care.

Managing publication and metrics under an OSI

Interviewees emphasized a critical need to ensure proper attribution for shared resources such that the MNI, its researchers, funders and donors benefit from their investments and efforts, while leveraging the advantages of greater openness. As such, some noted that the development of specific metrics for sharing would be of great practical benefit, allowing investments in the generation of data and other resources to be readily captured for use in funding and academic advancement evaluations. Thus, securing the recognition of such metrics by funders and McGill management would support adoption of OS by MNI researchers, and perhaps in the broader community.

Further, some researchers spoke of the need to create an “institutional memory for the source” of shared resources. Thus, one emphasized the importance of tracking the use and diffusion of shared MNI data/resources. She/he suggested that tracking could be facilitated by requiring users to register in order to access data/resources. Additionally, another noted that this mechanism could also allow for ‘vetting’ of users. Likewise, it was said that metrics on data downloads, website visits etc. should be gathered. Finally, researchers noted that standard academic citation by users of shared resources remains an important metric for influence. Thus, they reiterated the importance of having an existing citable publication describing the resource.

“It’s one thing to share and be open, but I think it is important to track because you know, at least we have to go back to the institution... if not the PI, getting the impact that it had on the community - so acknowledging the author or if lacking of reference, you know, citing the website where the data originated from. It’s track-able, so we need to find the proper solution for that”.

- Researcher, imaging and bioinformatics

Most interviewees agreed that standard academic authorship criteria should apply to attribution of resources shared by MNI researchers. Significant intellectual contribution is required to merit authorship, but use without further intellectual engagement should be generally be recognized with an acknowledgement or citation. However, a few researchers suggested that depending on the type of data shared and its importance in a new study, making the sharer a co-author for a defined time period after the data is publicly shared may be appropriate.

In the case of pre-publication sharing, some interviewees suggested that users should be subject to a publication moratorium until the sharer feels that they have extracted all useful information, perhaps within an associated time limit on this ‘first right to publication’. A few raised the idea of an additional limited post-publication moratorium on openly shared data and algorithms to allow the sharer more time to extract value.

Points to consider:

- Overall, interviewees underlined the importance of developing policy to clearly delineate which resources can be accessed, the level of oversight that will be imposed and the attribution that is required.
- A few suggested introducing an OS framework in a step-wise fashion, i.e. ‘testing the waters’ by progressively advancing toward greater openness.

Open access publication

The majority of interviewees said they aim to publish in the highest impact factor journal possible, without considering whether these are open access. A number of researchers are concerned that publishing in open access journals is considerably more expensive, and that the impact factors of these journals tend to be low. Further, a few interviewees suggested that open access journals are not necessary to facilitate broad access because universities generally hold subscriptions to pay-for-view journals, and because they themselves always respond to requests for publications when asked. Conversely, a few researchers favoured open access

journals for their rapid publication timelines, and publication of non-traditional or innovative stories.

Points to consider:

- Some interviewees noted that they would be more willing to publish in open access journals if the MNI were willing to cover the costs of the increased publication fees.
- One researcher suggested that the MNI could develop an infrastructure for early and open access by creating a pre-print server similar to those used in physics research.

9) Infrastructure, financial, human resource and time allocation

Sharing requires resources

Interviewees repeatedly expressed the concern that the wider sharing of data and scientific materials would impose a significant additional burden on them. Many emphasized that researchers are typically over-committed, noting that they are currently sharing to the extent possible considering infrastructure, time, financial and human resource limitations.

Interviewees provided us with concrete examples relevant to their research setting. Many discussed the extensive resources required to respond to requests for reagents. Clinician-scientists highlighted the investment required to collect and process patient samples, to collect and organize appropriate meta-data, and to maintain cell-lines. Imaging and informatics researchers spoke of the extensive requirements of building, maintaining and updating IT infrastructure. They also underlined the enormous challenges in ensuring interoperability and harmonization. Many researchers highlighted the statistics, IT and bioinformatics expertise needed to organize and format datasets for submission to journals and to repositories, or to send to other researchers. Many noted that obtaining REB approval to share data or use data others have shared involves substantial time and effort. Finally, as an example of the maintenance requirements of even a small database, a few researchers mentioned an intra-institute reagent registry that became out-dated, because of the lack of time and human resources needed to update it.

The Neuro Biobank

Researchers raised specific concerns about the Neuro biobank. While some are enthusiastic about this initiative, others warned that the MNI does not have a sufficient appreciation of the

resources required to establish, maintain and administer broad sharing of an extensive, comprehensively-annotated repository. In particular, clinician-scientists noted the difficulty in finding enough time during clinical appointments to properly educate patients and obtain informed consent for participation in research. They noted that explaining the implications of OS, particularly if this entails the use of more complex consent forms (see Theme 11: Public trust and consent for more details) will be more time-intensive. Citing the resources required to acquire and organize clinical data, they reiterated that the biobank should be advanced in a pragmatic and project-specific manner, guided by defined research questions to avoid open-ended and ultimately useless data acquisition.

Infrastructure for sharing

Interviewees expressed near universal satisfaction with non-profit repositories like Addgene and Jackson Labs, which administer broad sharing of plasmids and mice respectively. Researchers favour these because both divest individual laboratories of the time and cost of distribution, and increase the accessibility of reagents through their databases. However, repositories currently exist for only a limited range of reagents. When asked, almost all researchers agreed that establishing MNI institutional infrastructure to support sharing of data and other scientific resources would greatly facilitate participation in an OSI. They mentioned the following elements in particular as being critical components of this infrastructure:

- A designated person or team with bioinformatics, statistics and data management expertise. Notably, one interviewee said individuals with such top-notch, cross-disciplinary expertise are rare and in high demand.
- A point person for education, legal expertise and guidance on operationalizing OS. For example, this person could consult on ways to manage sharing or to navigate licensing and IP issues, conduct patent landscapes/freedom-to-operate analyses, etc.
- IT infrastructure: researchers gave a variety of examples, including building additional IT infrastructure and hiring more staff.
- Some clinicians indicated that a centralized office and process for patient education about the OSI and for obtaining consent could reduce the time burden on clinicians (see Theme 11: Public trust and consent for more details).
- Most researchers implied or directly stated that increased and sustained funding will be required. Several noted that traditional funding cycles are too short to support the needs

of an OSI. In particular, clinician-scientists and imaging and bioinformatics researchers underlined the significant resource requirements of biobanking and IT infrastructure.

10) Human samples/data and REB approval

Interviewees who work with patient data underscored the importance of ensuring responsible management of patient samples and information for the success of the OSI. They stated that their key ethical concern was the need to protect patient confidentiality. Researchers highlighted the role of the Neuro Research Ethics Board (MNI REB) in facilitating sharing of patient-derived resources under an OSI.

Researchers' opinions about the Neuro REB were mixed. Many interviewees were satisfied with their interactions with the REB, saying that it is helpful in advancing researchers' sharing and collaboration activities. For example, several were surprised by the permissiveness of the REB toward consent forms that allow for future use of samples. Likewise, researchers involved with the Neuro biobank said they are currently working with the REB to develop detailed governance mechanisms and patient consent protocols to allow broad sharing. Thus, some researchers said that research ethics norms are evolving in step with science. However, a few interviewees – notably some in the 55 and older age demographic – stated that REB requirements are sometimes overly stringent and a barrier to efficient sharing of their own research data, or to making use of patient data shared by others through open Web repositories. They said the REB process is onerous, and some have delayed engaging with it because of the anticipated time and effort it will involve. One researcher in the imaging and bioinformatics area specified that REBs often lack the technical understanding to put the risks of data-sharing in context. Some researchers, especially those who use patient-derived data or materials but do not themselves interact with patients, were uncertain of what the REB would require to approve the sharing of patient-derived data through open Web repositories. Several interviewees said that streamlining REB processes would aid in the implementation an OSI.

There's Quebec ethics... Canadian ethics, there's global, other countries' ethics... often those ethical boards don't have a very strong technical background. ...they'll say "Where is your data stored? Is it accessible to other people?" Well, technically any server, any computer in the entire planet is technically accessible to other people - any good hacker can find his way into, or her way into that... but those little nuanced explanations sometimes fail the ethics board's smell test. There's a whole bureaucratic process that

has to go through different ethical committees... just filling out forms and all that stuff, like it takes a long time.

- Researcher, imaging and bioinformatics

Lack of harmonization may slow OS

Researchers said that lack of alignment in the ethos, processes and requirements of REBs and of relevant legislation across jurisdictions can slow efforts to collaborate and share data. For example, interviewees involved in multi-centre clinical trials pointed out that research protocols must be approved by the REBs of all the collaborating institutions, and that is a time-consuming process. They added that the need to approve OS-related specifications, such as deposition of data in open repositories and the ability to re-use sample in future studies, may further extend timelines. Clinical researchers noted that coordinating REB approvals across North American jurisdictions is a challenge, and that navigating the international context can be even more complex. Imaging and informatics researchers who exchange data across international jurisdictions emphasized similar points. Thus, a number of researchers called for greater harmonization of legal and REB requirements within Canada and beyond.

Points to consider:

- Our interviews suggest that some researchers are confused about the Neuro REB's requirements – some stated concerns that were directly contradicted by others.
- At least one interviewee expressed frustration about the lack of transparency and inability to question or seek recourse regarding REB decisions. This interviewee called for the implementation of an external review process for REB decisions.
- Some interviewees suggested that clarification and education about the REB review process and requirements under an OSI would be an important facilitator of the initiative.

Our interviewee from the Neuro REB, Dr. Eugene Bereza, acknowledged that obtaining REB approval for studies involving multicentre research with human participants can be time-consuming and overly complex. However, he emphasized that the philosophy of the Neuro REB is calibrated toward expediting good research, and that there is data to support the efficiency of its review process.

He outlined several reasons why the REB approval process in general can be complex: 1) there are many different and sometime contradictory regulations that apply: different institutions often interpret and apply them differently. The fact that there is no professional standard or certification for REB membership may contribute to the lack of alignment between different REBs; 2) there may be a tendency for REBs to fail to achieve the fine balance between ensuring that risks to patients are minimized and facilitating the rapid advance of valuable research for the greater good. These issues are currently a concern for the efficiency of collaborative research and broad data-sharing within Quebec, nationally and internationally. He speculated that navigating REB approval processes may be a significant hindrance to the advance of OS if the above concerns are not addressed soon. However, he/she noted that the Neuro REB is conscious of, and actively working to mitigate them.

Dr. Bereza acknowledged that researchers often have a poor understanding of REB process and guidelines, often attributing delays to the REB when in fact they are due to the complexity of existing frameworks or the failure of sponsors to respond to REB requests. However, he pointed out that the MNI is in the rare position of having two professional ethicists, who are available to consult on protocol development.

- Our analysis suggests that the Neuro REB is supportive in assisting researchers in advancing OS, in alignment with ethical principles.
- Dr. Bereza felt that simplifying and harmonizing regulation pertaining to human subject protection in research is justified, and would do much to facilitate OS. He suggested that a coordinated approach involving government lobbying by research leaders would likely be the most effective approach.
- At the level of the OSI, he suggested that developing clear guidelines, informational materials for researchers, collaborators and patients about OS, and templates for REB documents, for example consent forms that allow for re-contact for secondary uses would greatly streamline the REB approval and research process.

11) Public trust and consent

In general, interviewees conflated public and patient trust, and focused on patient consent issues rather than on how the general public would view an OSI. Clinician-scientists and those

working in clinical research underlined the importance of informed consent processes and the stipulations of the agreements in consent forms for a successful OSI.

Perceived patient acceptance of OS

Interviewees were divided regarding how they thought patients would react to the notion of OS, and thus whether consent for research participation would be easier or more difficult to obtain. The majority tended to agree that patients mostly act in self-interest: if OS is presented as a means of accelerating translation to clinical benefit, then they would be receptive. A few felt that patients would be indifferent to an OS framework at the MNI, while others said that some patients would not be comfortable with the idea of their samples or data being shared beyond the immediate researcher's network, or used in as-yet-undetermined research.

Informed consent processes

A number of researchers noted that the success of the OSI, and more specifically the proposed Neuro biobank, would be optimized if all patients are considered potential research participants. A few proposed that dedicated personnel in a dedicated physical location should carry out patient education and consent to ease demands on clinicians' time, and to ensure in-depth and consistent engagement. However, a few interviewees were very concerned about the context in which consent is obtained. They emphasized that the MNI must take great care to avoid creating the perception that patient consent to OS is a prerequisite to receiving high-quality clinical care. Thus, one interviewee strongly stated that only the clinicians and their teams should obtain consent from patients. This interviewee was concerned that an OSI that involves default soliciting of patients for open sharing risks the misperception that the MNI is using patients to further its OS agenda, rather than that prioritizing patient care. This would be a risk for both the clinical and research activities of the MNI.

"...if we get one press release saying "this is just a place where you give your body", we're dead. We're just dead as a clinically-centred institution, we're dead... they will just not come".

- Clinician-scientist

Others spoke of the importance of ensuring that the therapeutic relationship does not compromise consent. Several interviewees warned that patients will agree to anything their physician recommends.

Informed consent formulation

Many interviewees emphasized the importance of obtaining informed consent to ensure patients' preferences and dignity are protected, and to enable efficient research progress (in terms of both time and cost). Several emphasized that consent forms that allow for ongoing use of data and materials, without requiring re-contact and re-consent for each new study, will be essential to a successful OSI. They said that current consent forms will need updating to address the needs of an OSI. A few supported the notion of 'blanket' or broad consents allowing for use in as-yet-undefined research. Others doubted that patients and their families would agree to such openness. Instead they favoured step-wise consents with yes/no clauses to ensure patient preferences are properly addressed and to maximize participation.

Public trust and donors

Some researchers stated that an OSI is likely to enhance public trust of the MNI. They stated that this is because of the potential for increased transparency, greater sharing of resources with a focus on accelerated translation, and less IP, which may emphasize that the benefits of research are channelled to patients, rather than industry. A branding strategy that underlines these benefits of OS at the MNI could be advantageous, if the primacy of safeguarding patient confidentiality and dignity under OS are emphasized. A few interviewees said they felt that OS will be attractive to donors. Indeed, one interviewee noted that donors have already expressed marked interest in funding their OS activities.

Some interviewees felt that public engagement could be further advanced through some form of online interface, including listings of current clinical trials, the details of MNI research activities, lay summaries and access to results of research in which patients were involved. Some did not favour such initiatives. One cited a general lack of scientific literacy in the public, and said access to results might cause more harm than good if patients misunderstand findings related to their disease.

Points to consider:

- Most researchers said communication and engagement with patients was of paramount importance. They stated that patients who feel that professionals have answered their questions carefully and comprehensively are more likely to agree to participate in OS research at the MNI, and that this calibre of engagement is required to ensure truly informed consent.
- Many interviewees said that it is critical to emphasize the precautions that the MNI is taking to protect patient privacy when presenting OSI to patients and promoting it to the public.

Analysis of patterns in opinions and concerns by interviewee demographics

We next sought to develop an over-arching perspective on the differences and similarities in researchers' viewpoints, and to explore visually whether sub-groupings of opinion exist and whether these segregate along demographic lines. We also wished to identify the key issues that divide the MNI constituency. The aim of the following analyses is to corroborate findings already made through our qualitative work and to identify additional associations and relationships to further drive analysis. We note, however, that our analysis is limited by the fact that it is based on qualitative rather than quantitative data. Qualitative methods are designed to identify trends and explore the bases for social phenomena. Thus, the interviews we conducted were semi-structured, such that the points that interviewees raised varied according to the knowledge and interests of the individual and the course of the interview. For example, while one person may have raised an issue or expressed an opinion, another interviewee may not have mentioned it or may not have devoted equal time to it (Patton 1990). Thus, our findings need to be understood in this context, and considered as an heuristic and hypothesis-generating, rather than as definitive. (For more details see Appendix 1: Methods, Quantitative analysis of coding responses).

We categorized material from our thematic analysis to a subset we called 'key categories'. These represent a focused subset of researcher opinions, concerns, ideas and motivations that we considered most relevant to the development of an OS policy at the MNI (see Appendix 3 for the list of categories). We carried out multi-dimensional scaling (MDS) and cluster analyses to visualize relative similarities between researchers with respect to their opinions about these key categories. We coloured the results according to research area in Figure 1 and 2, and by

gender in Figure 3. Note that classification of the research area for all MNI researchers was provided by MNI management before interviewees were invited to the study – we used this classification in our analyses, or as self-reported by interviewees if this differed from the MNI’s list.

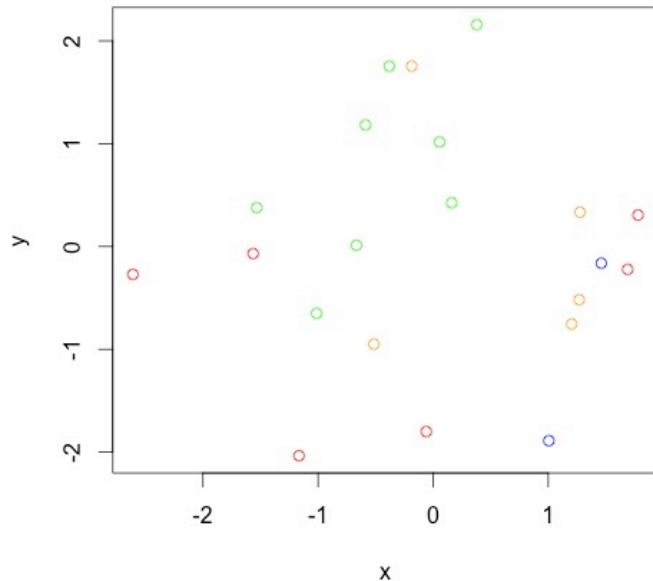


Figure 1: Results of multidimensional scaling (MDS) analysis projected in two dimensions based on similarity of interview content regarding key aspects of Open Science (OS). Each circle represents a single researcher with results coloured by research setting where green = cellular and molecular, orange = imaging and bioinformatics, blue = cognitive, and red = clinical and translational.

Figure 1 shows there was some tendency for interviewee opinions to cluster according to the research area in which they work. In particular, the MDS plot suggests that ‘cellular and molecular’ researchers (green) tended to hold similar opinions regarding key aspects of OS, as do the two researchers working in the ‘cognitive’ area (blue). Researchers from the ‘imaging and bioinformatics’ group (orange) show much looser clustering, while researchers working in ‘clinical and translational’ research (red) are found at opposite extreme edges of the plot, indicating dissimilarity in the responses with this group. To explore the reason for this spread we sub-classified the two ‘clinical and translational’ interviewees who are involved in clinical research, but do not run their own labs (purple). This suggests that the viewpoints of the researches conducting clinical research only are closely matched (purple), and diverge from the other clinical and translational researchers (Figure 2). Overall, these results suggest that

considerable heterogeneity exists with regard to key aspects of OS policy amongst the researcher constituency at the MNI, and even to some degree within research areas.

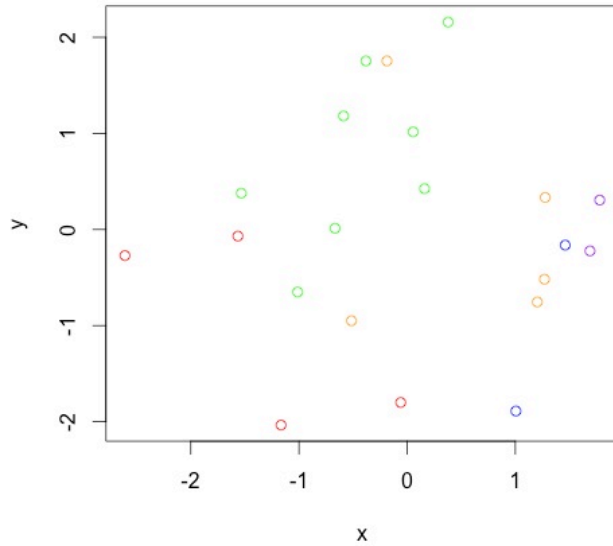


Figure 2: Results of MDS analysis (two dimensions) based on similarity of interview content regarding key aspects of OS, coloured by research setting where green = cellular and molecular, orange = imaging and bioinformatics, blue = cognitive, red = clinical and translational and purple = clinical only.

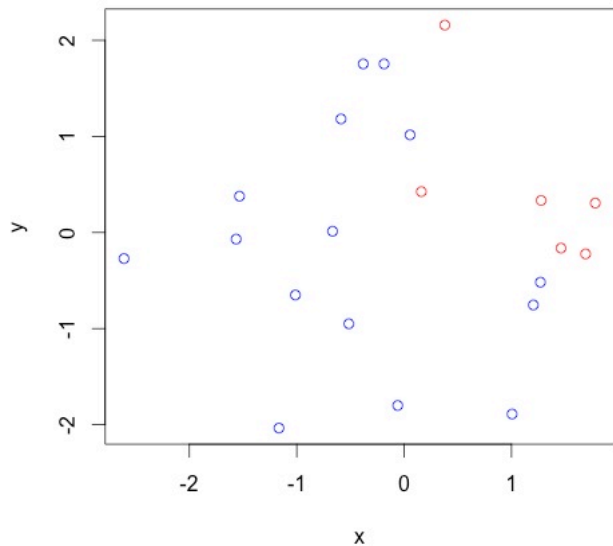


Figure 3: Results of MDS analysis (two dimensions) based on similarity of interview content regarding key aspects of OS, coloured by gender where red = female and blue = male.

To further explore how demographics map onto differences in researcher opinion, we coloured the same MDS projection by gender (Figure 3). Female researchers (red) cluster at one end of

the two-dimensional space, indicating that men and women tend to raise different concerns. This is especially interesting given that we interviewed at least one woman from each research setting, and that these women were diverse in terms of age, training and years in practice. Colouring the MDS projection by other demographic factors such as age, training and years in practice did not reveal any clustering (see Appendix 4). Figures 1, 2 and 3 reflect coding at our list of ‘key categories’. Similar analyses carried out with a larger group of categories from our thematic analysis (i.e. what we designated the ‘substantive categories’) showed similar patterns. See Appendix 4 for these and other supplementary figures.

Relative ideological positions of interviewees regarding OS

We also sought to classify researchers based on their perspectives on two central issues in the implementation of OS policy at the MNI: sharing of scientific resources, and intellectual property protections. To better explore the distribution of researchers along these ideological spectra, we created two indices: (1) an ‘opinion on resource sharing’ index to indicate the expressed comfort level of individual researchers toward sharing data and resources and (2) an ‘opinion on IP protections’ index to indicate how comfortable researchers are with forgoing intellectual property protections. For detailed methodology on how we developed these indices, refer to Appendix 1.

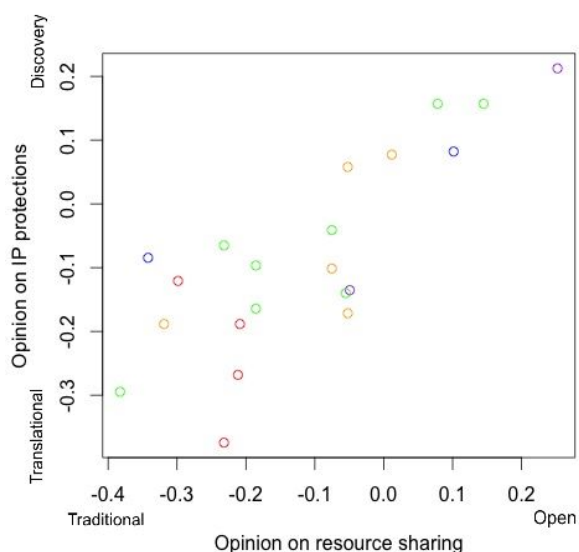


Figure 4: The relative position of interviewees by their ‘opinion on resource sharing’ and ‘opinion on IP protections’ index scores (based on our thematic coding). Each circle represents a single researcher with results coloured by research setting where green = cellular and molecular, orange = imaging and bioinformatics, blue = cognitive, red = clinical and translational, purple = clinical only. $P < 0.001$, $R^2 = 0.62$, confirmed by linear model analysis).

A high 'opinion on IP protections' score represents the perspective of a group we designated as 'discovery' researchers. These were interviewees who said that their main research motivation is advancing scientific knowledge, with little interest in claiming IP or pursuing commercialization. Conversely, a low 'opinion on IP protections score' indicates 'translational' researchers, those who said they seek IP and want to be actively involved in translation of these inventions. A high 'opinion on data sharing' score revealed researchers who supported 'open' sharing of data and scientific resources, while a low score represented researchers who adhere to a more 'traditional' model of data sharing; that is, they favoured greater control of how and with whom their resources are shared (for example, only engaging in bilateral sharing of resources with colleagues at their discretion). Together, these two indices allowed us to examine the distribution of interviewees in a two-dimensional space defined by two major ideological spectra relevant to an OS policy. These results are shown in Figure 4.

A high 'opinion on IP protections' score suggests a willingness to forgo patents, while a high 'opinion on resource sharing' score indicates a willingness to openly share data and materials. Our analysis therefore suggests that interviewees found in the top right of this two dimensional space are most likely to embrace an OSI at the MNI, while those in the bottom left are most likely to be resistant to such change. Figure 4 also shows a significant positive correlation between an individual researcher's 'opinion on resource sharing' and 'opinion on IP protections' index scores (confirmed by linear model analysis to $p < 0.001$, R^2 of 0.62). Thus, this analysis suggests that researchers who expressed concerns about open sharing of scientific resources often also said they favour pursuing and being the beneficiary of intellectual property rights on their outputs. Our analysis shows some clustering according to research area. For example, the two clinician-scientists who do not have labs of their own and carry out clinical research only are both situated at the top right of the plot, indicating that they favoured greater openness and had little interest in IP (purple). Conversely, the other 'clinical and translational' researchers are located closer to the bottom left of the plot, indicating their desire to have greater control over the diffusion of their scientific resources and intellectual property. 'Imaging and bioinformatics' researchers show some cohesion in their opinions, but interestingly they are not found at the leading edge of openness, according to this analysis.

We sought to further explore the index values of researchers by using ANOVA models to examine the influence of demographic factors including age, research setting, gender, job title,

and type of training. In predicting the 'opinion on resource sharing' index, only gender was a significant factor ($p=0.0012$); women had significantly higher values than men. Gender remained a significant factor in predicting the 'opinion on IP protections' index ($p=0.0017$), along with training ($p=0.046$), and age (marginally significant, $p=0.091$). Again women had higher values than men, PhDs had generally higher scores compared to MD/PhDs and MDs, and researchers 40 and under had higher scores than those over 40.

Cluster analyses

Our two indices suggested that research setting is a poor predictor of researchers' 'willingness to move towards OS'; hence, we next used a clustering algorithm to identify groups of ideologically similar researchers within the MNI constituency. These results are shown in Figure 5. The hierarchical cluster algorithm classified researchers based on their 'opinion on resource sharing' and 'opinion on IP protections' scores into one of two major groups and five distinct subgroups.

Major group I contains all the researchers with 'opinion on resource sharing' index scores of less than -0.1 (Figure 6). This group contains those most opposed to the open sharing of data and materials. Subgroup IA represents the strongest opponents of OS science; namely, those who prefer to keep their data and resources closed and to seek IP on their discoveries. Subgroup IB represents those who expressed more willingness to move towards open sharing of resources than subgroup IA, but still had a clear desire to pursue IP. Subgroup IC overall expressed less desire to patent, but still want considerable control over access to their data and materials; this suggests alignment with a more traditional academic research ethos. Major group II contains researchers most willing to be open with their scientific resources. Group IIA represents the most extreme sharers with the highest score in both indexes, while group IIB contains more moderate sharers and those with more interest in commercializing their discoveries. To provide further insight into the priorities of each subgroup we list the categories from our thematic analysis that underlie these groupings. Table 4 shows the categories that were over-represented in each subgroup.

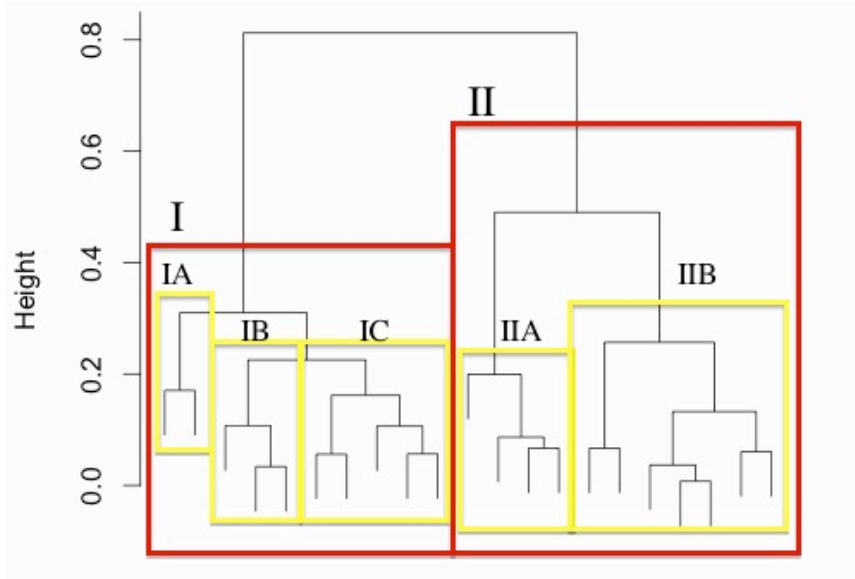


Figure 5: Hierarchical clustering of interviewees based on their 'opinion on resource sharing' and 'opinion on IP protections' index scores.

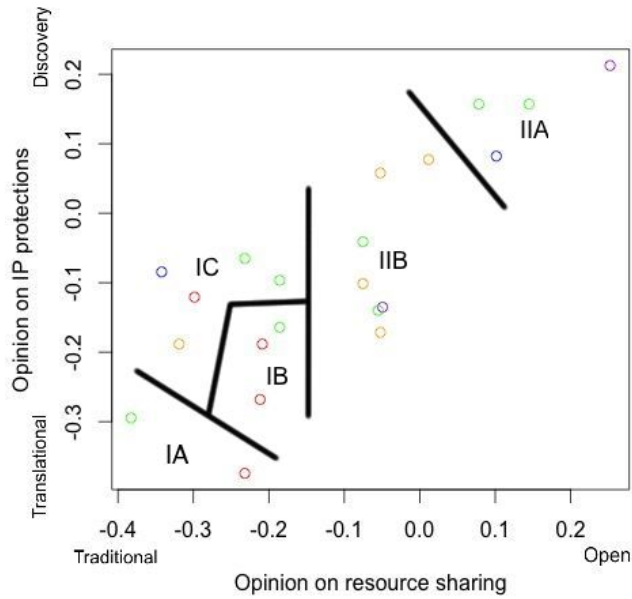


Figure 6: MDS plot indicating the subgroups identified by the hierarchical cluster analysis in Figure 5.

Table 5: Examples of categories most mentioned by each of the groupings shown in Figure 6. To provide more context for these categories, each one is listed with the higher level thematic coding category derived from our analysis.

IA	IB	IC	IIA	IIB
Researcher opinion on IP – IP is essential to encourage private investment in research	Changing paradigms in academic research and science in general – greater openness is a reality of contemporary research – current MNI policy is lagging behind	Disincentive to sharing or OS – concerns around privacy and confidentiality of participants	Infrastructure support for sharing – bioinformaticians and support staff	Issues around REB approvals – lack of understanding of REB approval processes
Collaborations – Collaborations with industry are needed to fund research translation	Researcher opinion on IP – IP is essential to encourage private investment in research	Motivations for sharing or OS – academic or ethical sharing ethos	Use of others’ patented research tools is not required for research	Motivations for sharing or OS – OS leads to more publications
Disincentives to share – Pre-publication sharing/ sharing of unvalidated findings may have adverse social impact through poor quality data and resulting research	Collaborations – Collaborations with industry needed to fund research translation	Traditional authorship standards persist for sharing research results and resources	Motivations for sharing or open science – ‘its the right thing to do’/‘Mother’s voice’/childhood lessons	Governance currently in place or seen as necessary in the context of an OSI – decisions around data sharing
Disincentives to share – risk of shared data being used out of context/ misused and wasting other researchers’ time	IP – researcher uses IP to protect their own research	Solutions – access to shared resources should be controlled	Expressed positive perspective on OS	Disincentives to sharing or OS – infrastructure is currently insufficient for sharing
Need for flexibility/ ability to opt-out of an OSI – academic freedom or PI personal freedom to choose	Need for flexibility/ ability to opt-out of an OSI – mandatory open access policy is a bad idea	The need for flexibility/ ability to opt-out of an OSI – academic freedom or PI personal freedom to choose	Solutions – Importance of clear goals for OSI	Motivations to collaborate with the MNI – OS/ biobank will attract collaborators through offering access to unique and comprehensive datasets

Separator categories

Finally, we sought to identify ‘separator’ categories to more effectively characterize important divisions within the MNI constituency with respect to OS policy. We did this iteratively by calculating an ‘explanatory value’ for each of the ‘substantive categories’ described above. The

explanatory value is a measure of how well the coding of a given category can predict the coding of all other categories across the interviews. We examined the categories with the highest explanatory value to validate and extend our qualitative understanding of the differences in opinion characterizing the constituencies at the MNI.

Results of our separator analysis indicated that the categories related to patenting and commercial development of scientific discoveries were the best predictors of interview coding. This corresponds with our findings in the previous section, where researchers in major group I (who expressed the desire to greatly restrict the flow of their resources and maintain ownership rights on discoveries), or major group II (who expressed willingness to share resources openly and forgo IP). Of the top five separator categories, three were opinions most associated with group II (i.e. researcher opinion on IP – IP is excessive red tape/slow research progress; lack of interest in commercialization; and researcher opinion on IP – IP is of negligible value). Conversely, two were opinions typical of group I (disincentives to sharing – researcher is reluctant to share because of their investment in creating resources and; researcher uses IP to protect their research). The ‘researcher uses IP to protect their own research’ category simply coded whether or not the researcher had claimed IP on her/his research outputs, and the fact that it was a significant predictor of interview coding suggests that researchers who hold IP have different priorities and concerns than the rest of the MNI in the context of an OSI.

REFERENCES

- Birney E et al.: Prepublication data sharing. *Nature*. 2009, **461**(7261), 168-170.
- Borg, I and Groenen, P: *Modern Multidimensional Scaling. Theory and Applications*. Springer.1997.
- Caulfield T, Harmon S, Joly Y: Open science versus commercialization: a modern research conflict. *Genome Medicine*. 2012, **4**(2): 17-27.
- Chesler E, Baker E: The importance of open-source integrative genomics to drug discovery. *Current opinion in drug discovery & development*. 2010, **13**(3): 310-316.
- Dove E, Özdemir V, Joly Y: Harnessing Omics Sciences, Population Databases, and Open Innovation Models for Theranostics-Guided Drug Discovery and Development. *Drug Development Research*. 2012, **73**(7): 439-446.
- Dyke S, Hubbard T: Developing and implementing an institute-wide data sharing policy. *Genome Medicine*. 2011, **3**(9), 60-67.
- Gower, JC: Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*. 1966, **53**(3-4), 325-338.
- Institute of Medicine (IOM). *Sharing clinical trial data: Maximizing benefits, minimizing risk*. Washington, DC: The National Academies Press. 2015.

<http://www.nap.edu/catalog/18998/sharing-clinical-trial-data-maximizing-benefits-minimizing-risk>

- Joly Y: Open biotechnology, licenses needed. *Nature*. 2010, **28**(5), 417-419.
- Mardia, KV: Some properties of classical multi-dimensional scaling. *Communications in Statistics-Theory and Methods*. 1978, **7**(13), 1233-1241.
- Marks L: TransCelerate: a global collaboration across biopharmaceutical R&D to accelerate and simplify new medicines development. *Journal of Medicines Development Sciences*. 2015, **1**(1), 14–19. <http://dx.doi.org/10.18063/JMDS.2015.01.010>.
- Morgan Jones M, Castle-Clarke S, Brooker D, Nason E, Huzair F, Chataway J: The Structural Genomics Consortium, A knowledge platform for drug discovery, A summary. Rand Corporation. 2014, http://www.rand.org/pubs/research_reports/RR512z1
- Murtagh, F: A survey of algorithms for contiguity-constrained clustering and related problems. *The computer journal*. 1985, **28**(1), 82-88.
- R Core Team: R, A language and environment for statistical computing. R Foundation for Statistical Computing. 2014, <http://www.R-project.org/>.
- Rodriguez H et al.: Recommendations from the 2008 international summit on proteomics data release and sharing policy, The Amsterdam principles. *Journal of Proteome Research*. 2009, **8**(7): 3689-3692.
- Organization for Economic Co-operation and Development: Collaborative Mechanisms for Intellectual Property Management in the Life Sciences. OECD Publishing. 2011.
- Patton M: Qualitative evaluation and research methods. Beverley Hills. CA: Sage. 1990.

APPENDIX 1: Methods

For this study we used qualitative research methods: in-depth semi-structured interviews, followed by thematic data analysis.

Study sample and research design

We used a purposeful sample strategy to select participants. The MNI provided us with a list of key MNI research areas and principal investigators which we used as our guide to selectively invite research participants, while aiming as much as possible for balanced representation across gender and research area.

Data collection

We conducted the research interviews in face-to-face meetings at the MNI in locations of the interviewees' choosing in June through September 2015. These were, except when no recording was authorized by the interviewee, digitally recorded and transcribed verbatim by our team of research assistants. Interviews lasted between 30 and 90 minutes. We developed the interview guides based on our review of the Open Science literature and knowledge of the MNI's research focus (see Appendix 2 for the Interview Guide) with the assistance of MNI management. The interview process was iterative, in that we fed key issues raised by interviewees into subsequent interviews.

During the interview process we also collected demographic information from our research participants. Where required and where it was possible, these data were verified through internet searches. The information collected included: researcher age; number of years of research experience; gender; job title; research setting; and type of training. Researcher age was recorded in one of three categories: 21-40, 41-55 or over 55. Years of experience (excluding time spent as a trainee, such as a PhD candidate) was recorded as: 1-5, 6-10, 11-20 or over 20. Job title captured whether researchers were full professors, assistant professors, post-doctoral researchers or research managers. Research setting was based on four categories provided to us by the MNI management: imaging and bioinformatics, cognitive, cellular and molecular, and clinical and translational. After the data was collected we further divided the latter category into interviewees that said they do clinical research only, and thus do not maintain their own research lab (called 'clinical only'). Type of training indicated whether the

interviewee held an MD and/or PhD. These demographic profiles were used to examine how ideas around the open science policy varied between groups, so as to identify constituencies in need of special attention in the development of an open science policy.

Qualitative thematic analysis

We have used qualitative thematic analysis techniques to analyze the data we obtained (Braun and Clarke, 2006). This process consists of 7 key phases: (1) familiarization—in which we read interview data in-depth multiple times; (2) generation of initial codes and application of codes to the dataset — or indexing — in which we identify and tag pieces of data (passages of text) relating to a common theme or idea relevant to the research question; (3) searching for and verification of themes across the entire dataset; (4) identification of relationships between codes and of distinct differences between subgroups of ideas; (5) definition and naming of themes; (6) re-reading of the interviews and modifying codes based on emerging themes; and finally (7) mapping and interpretation of the overall narrative identified from the data. We used NVivo 10 software (QSR International) to organize this process. Two RAs, Kendra Levasseur and Liam Harris, Sarah Ali-Khan and Richard Gold conducted and/or supervised the data analysis.

Development of analytical categories

During the thematic analysis process we developed several analytic categories into which we coded the content of our interviews. These categories capture what we defined as substantive, auxiliary and key data or issues. Substantive categories capture interviewees' opinions, concerns, ideas and motivations regarding Open Science, while auxiliary categories capture contextual information that we used to further our understanding of the substantive categories, such as the types of resources or stakeholders that interviewees mentioned as relevant. For example, where a researcher expressed concern about sharing iPS cells because of the large investment required to generate them, the text would be coded into a substantive category (“disincentives to sharing or open science – researcher time and money invested in resource creation”) and an auxiliary category (“types of resources – iPS cells and other cell lines”) to give contextual information.

We developed substantive categories to capture motivations for participating in an open science framework, including the potential to take advantage of collective problem solving, to establish new and diverse collaborations and to increase the quality of research and patient care.

Likewise, another set of substantive categories recorded major disincentives for participating in an open science framework, including fear of losing advantage to other researchers, insufficient infrastructure for sharing and fear of losing out on private investment. Other categories indexed solutions to overcome these disincentives that were mentioned by interviewees, including infrastructure support for sharing, metrics for measuring and reporting openness and controlled access mechanisms in the open dissemination of resources. We classified interviewees' opinions and interactions with intellectual property into categories such as distaste for ownership, necessary to spur investment and excellence over patents. The remaining higher order substantive categories deal with specific issues including changing academic paradigms, public/patient interest in an open science policy and the timeline for sharing academic resources.

After coding the interviews and verifying this coding across the dataset, we further distinguished a set of subcategories called 'key categories' for use in our multi-dimensional scaling and cluster analyses (see description below). The key categories represent a more focused subset of interviewees' opinions, concerns, ideas and motivations for the development of an open science policy at the MNI. These subcategories were largely drawn from the following substantive categories: (1) disincentives to open science, (2) motivations for open science, (3) researcher opinion on IP, as well as other frequently mentioned and/or divisive issues including: the need for investment in research translation; the need for a metric of openness; and the need for infrastructure support. For a full list of the categories, classifications and definitions developed through our analysis please refer to Appendix 3.

Inter-coder reliability

The research team undertook measures to ensure that the coding of the data was a reliable representation of interview content and that the coding framework was consistently applied across interview transcripts. The co-investigator and the two research assistants familiarized themselves with the breadth of interview content, and then collaboratively developed a preliminary coding framework. Throughout the interview coding process coders discussed changes to the code structure with each other to ensure uniformity, and the entire research team met on a weekly basis to discuss emerging findings. The interview coders also undertook a formal inter-coder reliability analysis using NVivo 10 software (QSR International), by both coding the same interview and measuring agreement between the two copies. Most inter-coder percentage agreement at each category varied from 80 to 100%, revealing only minimal

disagreement between coders. Categories at which agreement was lower were among the 'auxiliary' group. When we inspected discrepancies we observed that this was due to minor differences in the length of the text coders had selected for auxiliary categories. These discrepancies had minimal effect on our MDS, cluster and other analyses because the length of sections coded was irrelevant. These analyses were binary, therefore they were based on whether or not an interviewee had mentioned a particular issue (see section directly below).

Quantitative analysis of coding responses

Quantification of interview data

In order to perform quantitative analysis on our dataset, we needed to transform our qualitative interviews into quantitative data. We did this using NVivo's matrix coding feature, generating a binary matrix indicating the presence/absence of a given code in each of our interviews. We performed this analysis separately for: (1) the full list of substantive categories which were mentioned in, at minimum, 3 interviews, and (2) the key categories (both described in the 'development of analytical categories' section, above), which represented issues we determined to be the most controversial and relevant to the development of an open science policy at the MNI. NVivo's matrix coding feature also allowed us to explore the intersection between auxiliary categories and substantive categories by creating matrices coding the number of interviews in which the coding of the substantive and auxiliary categories overlapped. This was done twice, once examining the intersection between resource type and substantive categories, and again looking at the intersection between the type of stakeholders discussed and the substantive categories.

It is important to note that this procedure was meant to produce a quantitative data set that would corroborate findings already made qualitatively by the research team and to identify additional associations and relationships to further drive our qualitative analysis. We cannot draw conclusions purely from the quantitative data set for two important reasons. (1) Since we conducted semi-structured interviews the questions we asked varied over the course of the research process and according to the interests and knowledge of each particular interviewee. A different set of questions asked could have led to different issues raised in the interview even with the same interviewee, meaning that the interviews are not perfectly comparable. (2) In coding whether categories were mentioned by researchers as a binary we lose information about how often the issue was mentioned in a given interview and the enthusiasm (or lack

thereof) with which the point was made by that interviewee. That said, the coding was structured with quantitative analysis in mind, such that there were separate categories for interviewees who brought up an issue as a genuine area of concern and, where relevant, those who brought up an issue only to dismiss it as non-critical in terms of their participation in an open science policy. This analysis remained, therefore, useful in identifying and confirming areas of agreement and disagreement across the various research activities at the MNI.

Multidimensional scaling analysis (reported in Figures 1- 4)

We used classical multidimensional scaling analysis, sometimes called principal components analysis (Gower 1966), to visually represent the differences in interview coding across our interviewees. We started with the substantive binary coding matrix described above in *Quantification of interview data* and calculated the matrix of Euclidian distances between interviews in terms of their coding using the *dist* function from the *stats* package in R (Borg and Groenen 1997). We then used this distance matrix and the *cmdscale* function, also from the *stats* package in R, to generate two-dimensional coordinates that approximate the multidimensional distance from the original binary coding matrix (as in Mardia 1978). The two-dimensional coordinates were plotted and demographic patterns were assessed visually using colour coding in R.

Development of 'ideological indices' (reported in Figure 4)

Many of our analyses treat each code category as a separate dimension of a multidimensional analysis exploring the similarity of interviewees. However, many of the categories we defined seemed to be linked and we hypothesized that they are indicators of broader ideological positions. In order to better explore the distribution of researchers along the ideological spectra most relevant to the development of an OS policy at the MNI we created two indices: (1) a 'opinion on resource sharing' index to capture how comfortable researchers are with sharing data and resources and (2) an 'opinion on IP protections' index to capture how comfortable researchers are with forgoing intellectual property protections. We did this by using a plus/minus coding system, whereby we identified categories from our thematic coding indicative of the interview's alignment with one of the two ends of each index spectrum. In each case the positive end of the spectrum corresponds to the position most consistent with aspirational OS (i.e. maximizing openness regarding the sharing of resources, and forgoing IP on one's own research output and avoiding use of resources subject to the restrictive IP rights of others).

Then, for each researcher, we calculated the coding density at each end of the index spectrum and took the difference between these densities to obtain the index score for that researcher. Examples are mentioned below but detailed category lists can be found in Appendix 3.

The 'opinion on resource sharing' index

Opinions that indicated a willingness to openly share data and resources, such as a desire to move towards collective problem solving or to seek out new collaborators under an OS framework has the effect of pushing those who expressed these opinions toward the positive end of the spectrum. Conversely, opinions such as a concern that a requirement to openly share data would hinder the ability to establish collaborations, lead to researchers losing their competitive advantage, or lead to researchers to lose out on financial rewards had the effect of pushing interviewees who expressed these opinions to the negative end of the spectrum. Thus, it can be said that interviewees who are situated high on the 'opinion on resource sharing' index tend to be more 'open' sharers, favouring lower barriers to the diffusion of scientific data and materials within and outside the MNI, while interviewees with low 'opinion on resource sharing' index scores reflect more cautious sharers, who carry out more limited sharing or 'traditional'-style sharing i.e. bilaterally with colleagues on a case-by-case basis.

The 'opinion on IP protections' index

We included interviewee opinions such as 'lack of interest in commercializing research', 'distaste for ownership' and when a researcher stated that they did 'not hold IP' as indicators of the positive end of this spectrum (i.e. greater willingness to forgo IP). Conversely, categories such as the 'belief that IP is necessary to spur industry investment', 'stated 'experience commercializing research' and a researcher's stating that they 'hold IP' as indicators of the negative end of this spectrum. We called interviewees that scored high 'opinion on IP protections' index scores 'discovery' researchers, suggesting that they have little interest in pursuing the commercialization their research themselves. We called those interviewees who scored low 'opinion on IP protections' index scorers 'translational' researchers. These interviewees expressed interest in maintaining ownership and control over their research outputs.

Together these two indices allow us to examine the distribution of interviewees in a two-dimensional space defined by the two major ideological spectra relevant to an OS policy. This

two-dimensional plot was generated in R (R Core Team). We coloured these the patterns along demographic lines. Correlation between the two indices was assessed using the *lm* function in R. The same function was used to create a ANOVA model where demographic indicators: age, research setting, gender and job type were used as predictors of the ‘opinion on resource sharing’ index and the ‘opinion on IP protections’ index.

Hierarchical cluster analysis (reported in Figure 5)

To complement multidimensional scaling analysis, above, in the exploration of the similarity of interviewees with respect to their ideas and concerns around an OS policy, we used hierarchical cluster analysis. This method is useful because, instead of simply creating a visualization of relative similarity between all interviewees, it explicitly groups interviewees together based on that similarity in a dendrogram structure. As in multidimensional scaling analysis we first calculated the distance matrix using the *dist* function from the *stats* package in R. We then fed this matrix through the *hclust* function, also in the *stats* package, based on FORTRAN code by Murtagh (1985). We chose the ‘complete linkage’ algorithm, which starts with each interviewee in their own cluster and then iteratively joins the most similar clusters. The output was then plotted as a dendrogram representing the similarity between interviewees based on the coding matrix. This was done separately for the more complete ‘substantive category’ list and the ‘key category’ list as described above. A separate distance matrix was generated from the ‘opinion on resource sharing’ index and ‘opinion on IP protections’ index values and used for hierarchical clustering of interviewees based on their openness and opinion of IP. We then labelled the node tips to visually assess patterns in demographic indicators.

Identifying priority issues by constituency (reported in Table 4)

In order for us to associate particular categories with constituencies by differences in demographic information, we calculated the difference between the mean number of interviewees of a given constituency coded at each substantive category and the mean number of other interviewees coded at each category. This value therefore represented the increase in coding prevalence in the constituent group compared to the rest of the sample, where 0 indicated no difference in coding prevalence, positive values indicated higher coding prevalence in the constituent group and negative values indicated lower coding prevalence in the constituent group. We performed this analysis for constituent groups defined by each of the demographic categories, as well as for constituencies defined using *hierarchical cluster analysis*,

above. We then examined the categories that had the highest relative prevalence for each constituency in order to inform our qualitative understanding of the priorities and special needs of each constituent group.

Identifying separator categories

In order to better identify important divisions within the MNI constituency with respect to the OS policy, we undertook to identify ‘separator’ categories. We did this by iteratively calculating an ‘explanatory value’ for each of the categories in the coding matrix. The explanatory value is a measure of how well the coding of a given category can predict the coding of all the other categories across the interviews. First, we sorted the interviews into two groups based on whether or not they were coded at the category in question. Then the sum of the multidimensional distances between the members of each group with respect to all other categories was calculated using the *dist* function from the *stats* package in R. These were summed together and then divided by the total multidimensional distance between all interviews. This value was then subtracted from one such that a high explanatory value indicates that the coding of that category was a good predictor of the rest of the coding matrix. We examined the categories with the highest explanatory value to inform our qualitative understanding of the fault lines between constituencies at the MNI.

Auxiliary code analysis

We used the matrices coding the number of interviews in which the ‘auxiliary categories’ overlapped with each ‘substantive category’ to ascertain the key issues and ideas associated with different types of resources and stakeholders. After generating these matrices we extracted the top issues for each auxiliary category (resource or stakeholder type) and used them to enhance our qualitative understanding of the considerations of researchers using certain resource types or interacting with certain kinds of external stakeholders in the context of an OS policy at the MNI.

Ethics consideration

This study was approved by the McGill REB I. All interviewees provided written informed consent.

APPENDIX 2: INTERVIEW GUIDE

Project: 'Building a Framework for Open Science at the MNI – deriving guiding principles from key stakeholders: a qualitative research study'

Note that question probes or follow-ups are denoted with bullet points Demographic and contextual questions:

To begin, I would like to ask you a few questions about yourself:

1) First, what is your job title/position?

- Clinician-scientist
- Clinician
- Professor
- Research Associate
- Post-Doctoral researcher
- Graduate Student
- Research Manager

2) Within what type of setting do you work? (as many of the following that apply)

- Animal model laboratory
- Human subject laboratory
- Cellular imaging laboratory
- Brain imaging laboratory
- Behaviour laboratory
- Clinical research
- Other

3) Could you please describe your training?

- MD
- PhD
- MD/PhD
- Master's
- other

4) How many years have you been in practice?

- Less than 1 year
- 1 to 5 years
- 6 to 10 years
- 11 to 20 years
- More than 20 years

On Open Science in general

1. In your view, what is Open Science?

2. Have you had any experiences publicly sharing data, samples or research materials? What was the context? What was the outcome?

- What motivated your participation?
- What problems or challenges did you experience, or did you anticipate?
- What were the benefits or harms if any, to the research process and results?
-

3. Have you had any experiences with patents or other intellectual property rights held by others which affected your research? Have you had any experiences patenting or asserting other intellectual property rights over discoveries you have made?

- What motivated you to seek IP protection or to use materials or processes subject to IP protection?
- What problems or challenges did you experience, or did you anticipate?
- What were the benefits to the research process and results?

4. What do you consider to be the appropriate boundaries of Open Science?

- What scientific resources are you/would you be comfortable sharing?
- With what limitations or caveats? Why?

For example:

- Types of data: Primary and processed datasets and associated meta data; patient genetic information and associated medical records (drug treatment regimes, biochemical data, test results, behavioral information, EEGs etc); failed or negative experimental data.
- Scientific tools/materials: experimental protocols, reagents, cell-lines, animal models, primers/genetic sequences, antibodies, patient tissues samples and extracts etc)
- Research reports, algorithms, software code, patented inventions etc.
- At which stage of the research process would it be appropriate to share these items? (eg. immediately upon collection, after processing, after manuscript submission, upon first publication). What factors inform your answer?
- Under what circumstances would it be appropriate to forgo/avoid intellectual property protection (like patents)? Under what circumstances would you consider it inappropriate? Why? Human participants and an obligation to share

5. What considerations, benefits or limitations do you see to Open Science when this involves sharing of research participant or patient-participant data or samples? (if applicable)

- How might this affect your relationship with participants, and the rights and duties owed them through the informed consent process? How might this affect how you obtain informed consent from participants? How might this effect public trust in the MNI's research?
- How could this affect the ability to secure Research Ethics Board approval for studies with human participants?
- How could these challenges best be addressed? What solutions may exist? How might the benefits be maximized?

Research implications

6. Would an obligation to share data and other scientific resource publicly have technical or logistical implications for your research program – for example might it make it easier to manage and organize data and materials, or conversely might it raise challenges?

- What impact could these challenges or benefits have on how you conduct research? (eg. preparing data to fit the format/type requirements of existing repositories may require increased time and resource commitments or it may create an incentive to streamline the research process etc.)
- What challenges or benefits do you foresee in obtaining proper attribution for and recognition of your work product?
- How could these challenges best be addressed? What accommodations may be necessary? What could be done to maximize the benefits of sharing?
- To what extent do appropriate repositories/infrastructure for the deposition of the data, samples or materials you collect or generate exist?
- What kinds of institutional supports would facilitate or otherwise encourage you to more freely share data, samples and/or materials?

7. If Open Science at the MNI constitutes to some degree a ‘patent-free zone’ how might this benefit your research? Conversely, are there ways in which it could raise challenges?

- How might the obligation to forgo intellectual property on your discoveries change how you conduct your research?
- How might the need to avoid using reagents or other patented inventions to the extent that these are incorporated in your results (and thus limit sharing) affect your research?
- If there are challenges, how could they be addressed? What accommodations may be necessary? What could be done to maximize the benefits?

Research partnerships

8. How might an institutional mandate to share data, samples and materials play a role in establishing or maintaining partnerships with other publicly-funded entities?

- In what ways could this enhance these relationships?
- In what ways could this challenges these relationships?

9. How might an institutional mandate to share data, samples and materials play a role in establishing or maintaining partnerships with privately funded or for-profit third parties?

- In what ways could this enhance these relationships?
- In what ways could this challenges these relationships?
- If partners are concerned, what might the issues be?
- What may be the boundaries on sharing from the perspective of privately-funded third parties?

10. How might an institutional mandate not to invoke patents or other intellectual property protections on the products of research play a role in establishing or maintaining partnerships

with privately funded or for-profit third parties?

- In what ways could this enhance these relationships?
- In what ways could this challenges these relationships?
- If partners are concerned, what might the issues be?!
- What may be the boundaries of a commitment not to claim intellectual property protections from the perspective of privately-funded third parties?

Key priorities and additional factors

11. In your view what would be the key elements of an institutional Open Science policy at the MNI?

12. Is there any other important aspect that we have not yet discussed?

Thank you for participating!

In the case of any questions or concerns please contact the co-investigator:

Sarah Ali-Khan, PhD
Research Associate, Project Manager
Faculty of Law, McGill University
New Chancellor Day Hall
3644 Peel Street, Rm 515
Montreal, Quebec
H3A 1W9
+514 926 9482
sarah.ali-khan@mcgill.ca

APPENDIX 3: CODING AND ANALYTICAL CATEGORIES

Substantive Categories

We developed a list of analytical categories that we call ‘substantive categories’. These represent the set of analytical coding categories capturing the substantive material of the interviews: the interviewees’ opinions, experiences and issues relevant to an OSI at the MNI. This large category includes the subset we derived and called ‘key categories’ (see below for details)

Attribution:

- Reasons to need attribution for your work & interaction with funding committees
- What types of attribution are expected and actually occur for sharing resources now
- Traditional authorship standards persist for sharing research results and resources

Changing paradigms in academic research and science in general:

- MNI policy is lagging behind modern reality of openness
- Opinions about the MNI’s transitional period if they decide to move towards open science

Collaborations:

- Require possibility of patenting discovery
- Require secrecy about data generated
- Collaborations with industry needed to fund research translation

Commercialization of research output:

- Researchers’ experiences with commercialization
- Researchers’ lack of interest in commercializing their research themselves
- Increased pressure felt by researchers to commercialize their research
- Commercial entities would take advantage of openness and commercialize research outputs without permission or collaboration, to the detriment of researchers

Disincentives to sharing or open science:

- Commercial entities would take advantage of openness
- Concerns around privacy and confidentiality of participants
- Currently insufficient infrastructure for sharing
- Lack of interoperability between potential sharers
- Issues around REB approvals
 - o Ethics approval processes as a barrier
 - o Lack of understanding of REB approval processes
 - o Positive aspects of REB involvement in move towards OS
- OSI conflict with McGill Patent policy
- Losing competitive advantage by sharing too much
 - o Non-issue
- Losing out on financial rewards for discovery
- Losing private investment in research
 - o Non-issue

- Pre-publication sharing/sharing of unvalidated findings may have adverse social impact through poor quality data and resulting research
- Researcher time and money invested in resource creation
- Time and cost burden of distributing shared resources
- Risk of shared data being used out of context/misused and wasting other researchers' time

Effects of an OSI on public funding received by the MNI:

- Public funding will increase
- Public funding will decrease

Expressed positive perspective on open science

Governance mechanisms currently in place or seen as necessary in the context of an OSI:

- Decisions around data sharing
- Decisions around IP
- Decisions around use and sharing of valuable resources
- Decisions around use and sharing of depletable resources

Intellectual property (IP):

- Use of others' patented research tools
- Use of others' patented research tools is not required for research
- Researcher uses IP to protect their research
 - o Copyright
 - o Licenses
 - o Patents

Motivations for sharing or open science:

- Academic and ethical sharing ethos
- OSI to attract new students and trainees
- Efficient collective problem solving
- Increase research impact and dissemination
- Increased quality of care and research
- Increased stature of MNI
- More and different collaborations
- OS leads to more publications
- "Mother's voice"/childhood lessons
- Return output of publicly funded research and public resources to the public

Motivations to collaborate with the MNI:

- Unique expertise or knowledge
- Comprehensiveness of integrated datasets

Patient-public interest in open science:

- Interactions between an OSI and consent
 - o Decreased ease of consent
 - o Increased ease of consent
- Public interest in OS is high
- Public interest in OS is low

- Public trust of the MNI will increase
- Public trust of the MNI will decrease

Researcher opinion on IP:

- IP provides advantageous royalty fees
- Holding IP protections is a valuable addition to a CV
- Distaste for ownership as a public researcher
- Excellence in research is more important than patents
- Excessive red tape involved in seeking IP
- IP is essential to encourage private investment in research
- IP protections have negligible value
- Uncertainty about the value of IP protections

Solutions:

- Communication and engagement
- Access to shared resources should be controlled
- Education of stakeholders (patients and public or researchers)
 - o Researchers
- Need for flexibility or ability to opt-out of an OSI
 - o Academic freedom or PI personal freedom to choose
 - o Mandatory open access policy is a bad idea
 - o Need for different models of engagement
- Importance of clear goals for OSI
- Infrastructure support for sharing
 - o Biobanks and data repositories
 - o Bioinformaticians and support staff
 - o Creation of a patient web interface
 - o Creation of a patient web interface
 - Negative opinion
- Need for a metric or incentive or credit for openness
- Partnership with REB in OSI development and implementation
- Publication in open access journals
- Reasonable costs charged for open provision of resources

Timeline for sharing:

- Pre-publication sharing of research-associated resources
- Post-publication sharing of research-associated resources

Auxiliary Categories

Auxiliary categories represent the analytical coding categories that we used to cross-reference interviewees' opinions, experiences and issues. They include for example, information about the area of research, the reagent, the data-type or the stakeholders the interviewee was speaking about. For example, using these categories in our analysis allowed us to compare how interviewees' opinions about IP protections differed across types of data or with respect to different stakeholders.

Experiences with sharing:

- Data sharing
- Not yet sharing but willing
- Protocols
- Reagents
 - o Only post-publication
- Software, algorithms

Intellectual property:

- Researcher doesn't use IP
- Researcher uses IP to protect their research (licences, patents, copyright)

Past (every mention of anything in the past)

Stakeholders:

- Foundations and private donors
- Journals
- McGill University
- MNI
 - o CRU
 - o Neuro biobank
- Private investors
 - o Biotechnology
 - o Pharmaceutical
 - o Start-ups
- Public collaborators
- Public funding agencies
 - o CIHR
 - o FRQS
 - o NIH
- Public or patients
- Public repositories
 - o Addgene
 - o Biobanks
 - o EEG
 - o FlyBank
 - o Jackson Labs
 - o Parkinson's Biobank
 - o PPMI
 - o Quebec Parkinson's Registry
- REB
- SGC
 - o Aled Edwards
 - o SGC model
 - Uncertainty about merit
 - Unsure about usefulness
- Spin-off companies
- Young researchers

Types of resources:

- Data
 - o Brain imaging data
 - o Experimental data
 - o Negative data
 - o Genetic data
 - o Patient clinical data
 - o Proteomics datasets
- Knowledge
 - o Expertise
 - o Methodologies, protocols, task batteries
 - o Tacit
- Reagents
 - o Animal models
 - o Antibodies
 - o Assays
 - o Compounds
 - o iPS cell and other cell-lines
 - o plasmids
- Software and technology
 - o Algorithms
 - o Products and platforms
- Tissue samples

Key Categories

'Key categories' represent the subset of interviewees' opinions, concerns and motivations that we determined are the most relevant to the development of an OS policy at the MNI. To a large degree this material represents the most significant sources of disagreement or tension about the proposed shift to OS, and material that was the most emphasized by interviewees.

Collaborations:

- Require possibility of patenting discovery
- Collaborations with industry needed to fund research translation

Disincentive to sharing or OS:

- Commercial entities would take advantage of openness
- Concerns around privacy and confidentiality of participants
- Currently insufficient infrastructure for sharing
- Lack of interoperability between potential sharers
- Issues around REB approvals
 - o Ethics approval processes as a barrier
 - o Lack of understanding of REB processes
 - o Positive aspects of REB involvement in Open Science
- Losing competitive advantage by sharing too much
 - o Non-issue
- Losing out on financial rewards for discovery
- Losing private investment in research
- Losing private investment in research

- Non-issue
- Pre-publication sharing/ sharing of unvalidated findings may have adverse social impact through poor quality data and resulting research
- Researcher time and money invested in resource creation
- Time and cost burden of distributing shared resources
- Risk of shared data being used out of context/misused and wasting other researchers' time

Governance mechanisms currently in place or seen as necessary in the context of an OSI:

- Decisions around data sharing
- Decisions around IP

Researcher opinion on IP:

- IP provides advantageous royalty fees
- Holding IP protections is a valuable addition to a CV
- Distaste for ownership as a public researcher
- Excellence in research more important than patents
- Excessive red tape involved in seeking IP, or IP itself is excessive red tape
- IP is essential to encourage private investment in research
- IP protections have negligible value
- Uncertainty about the value of patents

Motivations for sharing or open science:

- Academic and ethical sharing ethos
- Efficient collective problem solving
- Increase research impact and dissemination
- Increased quality of care and research
- Increased stature of MNI
- More and different collaborations
- OS leads to more publications
- Return output of publicly funded research and public resources to the public

Solutions:

- Communication and engagement
- Need for flexibility or ability to opt-out of an OSI
 - Academic freedom or PI personal freedom to choose
 - Mandatory open access policy is a bad idea
 - Need for different models of engagement
- Importance of clear goals for OSI
- Infrastructure support for sharing
- Need for metric or incentive or credit for openness

Timeline for sharing:

- Pre-publication sharing of research-associated resources
- Post-publication sharing of research-associated resources

Positive perspective on open science

Gradient Categories

We used the gradient categories to develop the indices used in Figures 4, 5 and 6, as explained in the ‘Relative ideological positions of interviewees regarding OS’ section in the main text and in Appendix 1. These categories capture issues that we determined to be important and/or controversial with respect to the implementation of OS at the MNI, and we used these to inform the index values.

Opinion on Resource Sharing

Coding categories we classified as ‘traditional’ captured interviewee opinions that favour allowing more control of what is shared and when. For example, sharing bilaterally with colleagues on a case-by-case basis. Conversely, categories included at the ‘Open’ sharing end of the gradient tend toward favouring lower barriers to the diffusion of scientific data and materials within and outside the MNI.

Traditional	Open
Attribution: <ul style="list-style-type: none"> - Traditional authorship standards persist for sharing research results and resources 	Changing paradigms in academic research and science in general: <ul style="list-style-type: none"> - MNI policy is lagging behind modern reality of openness
Collaborations: <ul style="list-style-type: none"> - Require secrecy about data generated 	
Disincentive to sharing or OS: <ul style="list-style-type: none"> - Issues around REB approvals <ul style="list-style-type: none"> o Ethics approval processes as a barrier - Losing competitive advantage by sharing too much - Losing out on financial rewards for discovery - Losing private investment in research - Risk of shared data being used out of context/misused and wasting other researchers’ time 	Expressed positive perspective on OS
Motivations for sharing or open science: <ul style="list-style-type: none"> - Academic and ethical sharing ethos - Increase research impact and dissemination - More and different collaborations 	Motivations for sharing or OS: <ul style="list-style-type: none"> - Academic and ethical sharing ethos - OSI to attract new students and trainees - Efficient collective problem solving - Increase research impact and

<ul style="list-style-type: none"> - OS leads to more publications 	<p>dissemination</p> <ul style="list-style-type: none"> - Increased quality of care and research - Increased stature of MNI - More and different collaborations - OS leads to more publications - “Mother’s voice”/childhood lessons - Return output of publicly funded research and public resources to the public
<p>Solutions:</p> <ul style="list-style-type: none"> - Access to shared resources should be controlled - Need for a metric or incentive or credit for openness 	<p>Patient-public interest in OS:</p> <ul style="list-style-type: none"> - Interface between an OSI and consent <ul style="list-style-type: none"> o Increased ease of consent - Public trust of the MNI will increase
<p>Timeline for sharing:</p> <ul style="list-style-type: none"> - Post-publication sharing of research-associated resources 	<p>Solutions:</p> <ul style="list-style-type: none"> - Infrastructure support for sharing <ul style="list-style-type: none"> o Biobanks and data repositories o Bioinformaticians and support staff o Creation of a patient web interface o Centralized patient registry or patient web interface o Legal expertise

Opinion on Intellectual Property Protections

Coding categories we classified as ‘translational’ captured interviewee opinions that favour maintaining ownership over their discoveries through holding IP. Conversely, categories included at the ‘discovery’ end of the gradient captured the viewpoints indicative of disinterest in pursuing the commercialization of their research output themselves.

Translational	Discovery
<p>Collaborations:</p> <ul style="list-style-type: none"> - Collaborations with industry needed to fund research translation 	<p>Commercialization of research output:</p> <ul style="list-style-type: none"> - Researchers’ lack of interest in commercializing their research themselves
<p>Commercialization of research output:</p> <ul style="list-style-type: none"> - Researchers’ experiences with commercialization 	<p>Collaborations:</p> <ul style="list-style-type: none"> - Lose sole first author status - Lack of interest in dealing with pharmaceutical companies
<p>Disincentives to sharing or OS:</p>	<p>Disincentives to sharing or OS:</p>

<ul style="list-style-type: none"> - Commercial entities would take advantage of openness - Losing out on financial rewards for discovery - Losing private investment in research 	<ul style="list-style-type: none"> - Commercial entities would take advantage of openness <ul style="list-style-type: none"> o Non-issue
<p>Governance currently in place or seen as necessary in the context of an OSI:</p> <ul style="list-style-type: none"> - Decisions around IP 	<p>Governance currently in place or seen as necessary in the context of an OSI:</p> <ul style="list-style-type: none"> - Decisions around data sharing - Decisions around valuable resources
<p>IP:</p> <ul style="list-style-type: none"> - Researcher uses IP to protect their research (aggregated) 	<p>IP:</p> <ul style="list-style-type: none"> - Researcher doesn't use IP to protect their own research
<p>Motivations for sharing or OS:</p> <ul style="list-style-type: none"> - Efficient collective problem solving - Increase research impact and dissemination - Increased quality of care and research - More and different collaborations 	<p>Motivations for sharing or open science:</p> <ul style="list-style-type: none"> - Academic and ethical sharing ethos - OSI to attract new students and trainees - Efficient collective problem solving - Increase research impact and dissemination - Increased quality of care and research - Increased stature of MNI - More and different collaborations - OS leads to more publications - "Mother's voice"/childhood lessons - Return output of publicly funded research and public resources to the public
<p>Researcher opinion on IP:</p> <ul style="list-style-type: none"> - IP provides advantageous royalty fees - IP is essential to encourage private investment in research - Holding IP protections is a valuable addition to a CV 	<p>Researcher opinion on IP:</p> <ul style="list-style-type: none"> - Distaste for ownership as a public researcher - Excellence in research is more important than patents - Excessive red tape involved in seeking IP - IP protections have negligible value - Uncertainty about the value of IP protections - IP only for defensive purposes
<p>Solutions</p> <ul style="list-style-type: none"> - Importance of clear goals for OSI - Need for flexibility or ability to opt-out of an OSI <ul style="list-style-type: none"> o Academic freedom or PI personal freedom to 	

<p>choose</p> <ul style="list-style-type: none">○ Mandatory open access policy is a bad idea○ Need for different models of engagement	
--	--

APPENDIX 4: SUPPLEMENTARY FIGURES

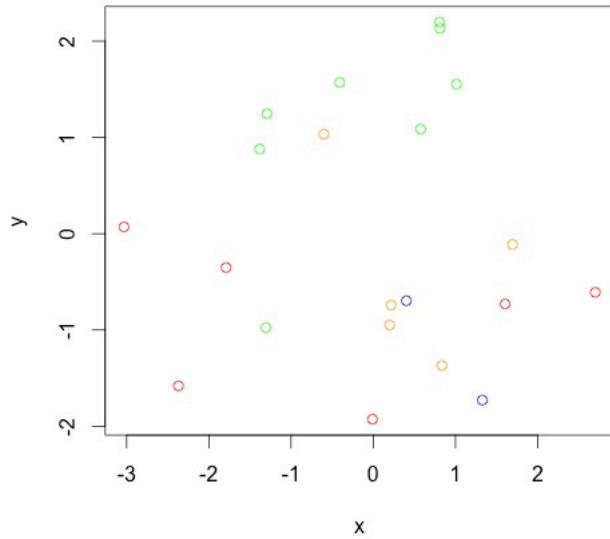


Figure S1: Results of multidimensional scaling (MDS) analysis projected in two dimensions based on similarity of interview content regarding all substantive aspects of Open Science (OS). Each circle represents a single researcher with results coloured by research setting where green = cellular and molecular, orange = imaging and bioinformatics, blue = cognitive, and red = clinical and translational.

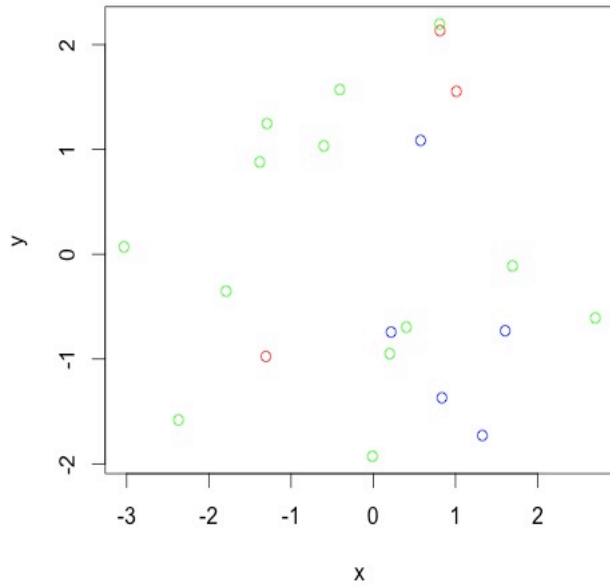


Figure S2: Results of multidimensional scaling (MDS) analysis projected in two dimensions based on similarity of interview content regarding all substantive aspects of Open Science (OS). Each circle represents a single researcher with results coloured by researcher age where red = 21-40, green = 41-55, and blue = 55+.

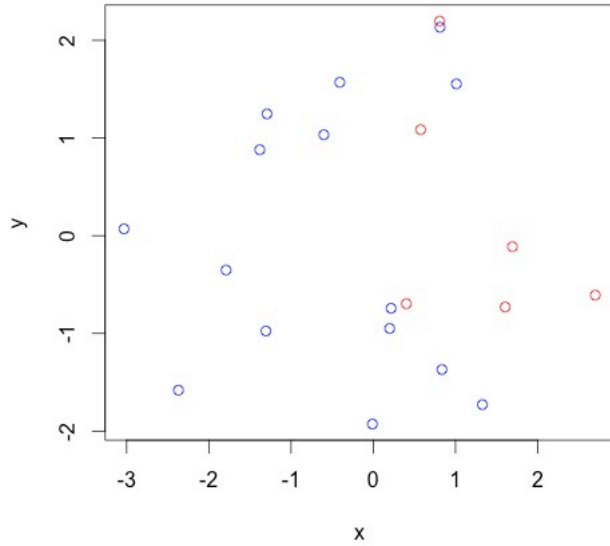


Figure S3: Results of multidimensional scaling (MDS) analysis projected in two dimensions based on similarity of interview content regarding all substantive aspects of Open Science (OS). Each circle represents a single researcher with results coloured by researcher gender where red = female and blue = male.

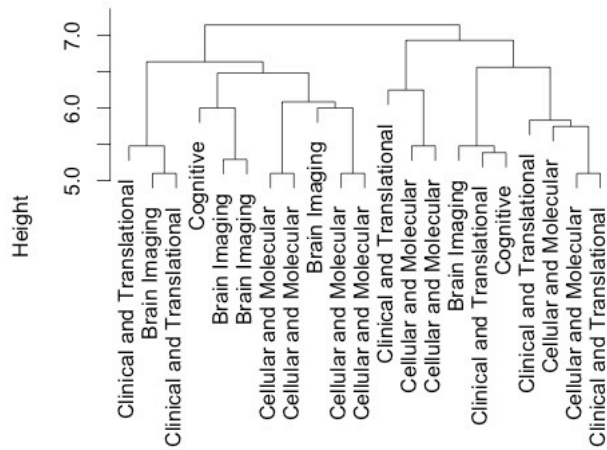


Figure S4: Hierarchical clustering of interviewees based on coding similarity at all substantive aspects of Open Science and labelled according to research setting.

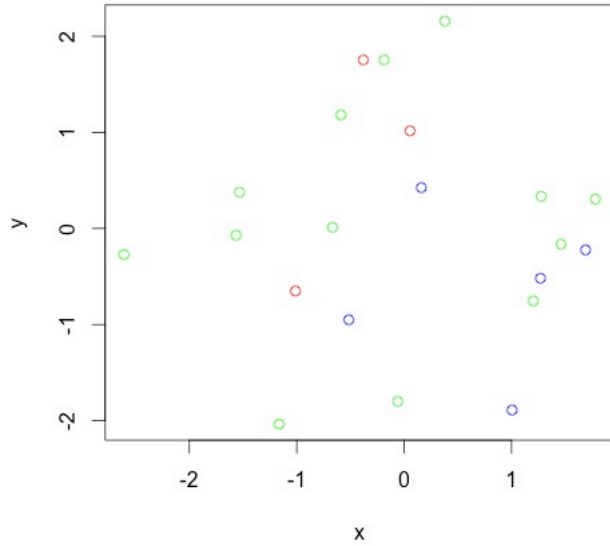


Figure S5: Results of multidimensional scaling (MDS) analysis projected in two dimensions based on similarity of interview content regarding key aspects of Open Science (OS). Each circle represents a single researcher with results coloured by researcher age where red = 21-40, green = 41-55, and blue = 55+.

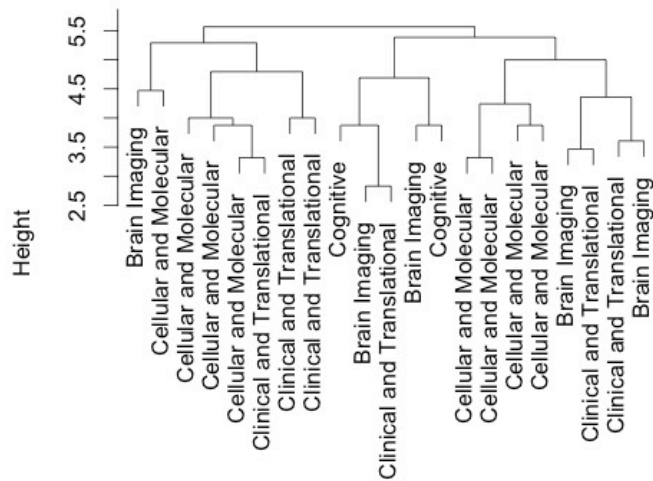


Figure S6: Hierarchical clustering of interviewees based on coding similarity at key aspects of Open Science and labelled according to research setting.